ISO/IEC JTC 1/SC 29 /WG3 N 00802

Date: 2023-01-20

ISO/IEC 23001-17:202X(E)

ISO/IEC JTC 1/SC 29/WG 3

Secretariat: JISC

Information technology — MPEG Systems technologies — Part17: Uncompressed video and images in ISO Base Media File Format

Document type: Draft International Standard

Document subtype:

Document language: E

Warning

This document is not an ISO International Standard. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an International Standard.

Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

Contents

Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents) or the IEC list of patent declarations received (see http://patents.iec.ch).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information Technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia, and hypermedia*.

A list of all parts in the ISO/IEC 23001 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Information technology — MPEG Systems technologies — Part 17: Uncompressed video and images in ISO Base Media File Format

1 Scope

This document defines how uncompressed 2D image and video data is carried in files in the family of standards based on the ISO base media file format. This includes but is not limited to monochromatic data, colour data, transparency (alpha) information and depth information.
The primary goal of this specification is to allow exchange of uncompressed video and image data while relying on the information set provided by the ISO base media file format, such as timing, colour space and sample aspect ratio to specify the interpretation and/or display of video and image data.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 14496-12, *Information technology — Coding of audio-visual objects — Part 12: ISO base media file format*

ISO/IEC 23008-12, Information technology – *High efficiency coding and media delivery in heterogeneous environments – Part 12: Image File Format* (HEIF)
ISO/IEC 23091-2, *Information technology —Coding-independent code points — Part 2: Video*
ISO/IEC 23002-3, *Information technology — MPEG video technologies — Part 3: Representation of auxiliary video and supplemental information*
IEEE 754, *IEEE Std 754™-2008, IEEE Standard for Floating-Point Arithmetic*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 14496-12 and the following apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at http://www.electropedia.org/

3.1 block

consecutive bytes within the sample data containing one or more component values for one or more pixels and possible padding

3.2 component

part of the image data representing a single channel (or dimension) of the image

Note to entry: In this specification, a component may describe visual information such as luminance or chroma, or other information usually not intended for direct display such as depth or transparency.

3.3 frame

two-dimensional rectangular array of pixels contained in the sample data

3.4 interleaving

pixel or component order within the sample data of component values

3.5 pixel

smallest element of an image, comprised of one or more components

3.6 row

horizonal line of pixels within a frame or a tile

3.7 sample data

payload of the media sample when the uncompressed frame is described by a media track, or payload of the item when the uncompressed frame is described by an image item

Note 1 to entry: Media sample as defined in ISO/IEC 14496-12.

Note 2 to entry: Image item as defined in ISO/IEC 23008-12.

3.8 tile

two-dimensional rectangular array of pixels within a frame

3.9 uncompressed frame

frame for which each value of each component is coded independently from any other component value in the same frame or any other frame

Note to entry: In this specification, the uncompressed term is used with some video formats applying sub-sampling of some components for the purpose of data reduction; however, data access to each individual component for such formats is still independent from other components or frames.

3.10 uncompressed image

single uncompressed frame stored as an image item

3.11 uncompressed video

sequence of one or more uncompressed frames

4 Uncompressed video and image formats

4.1 Overview

Uncompressed video formats may be stored in ISO base media files as media tracks or image items using a generic uncompressed video description defined in this specification.

Media tracks, media samples and image items may be associated with meta-data information using the various tools defined in ISO/IEC 14496-12, such as sample group descriptions, `MetaBox`, metadata tracks and sample auxiliary information. User-defined components may be used to carry per-pixel metadata, either in the same sample or item as the described pixels or in a separate track or item.

The tools defined in ISO/IEC 14496-12 and ISO/IEC 23008-12 should be used whenever applicable, namely to specify pixel aspect ratio, colour information, clean aperture, content light level, mirror and rotate properties or track header matrix, etc.

An uncompressed video media sample or item consists of one uncompressed frame. Each uncompressed frame is organized as a set of one or more rectangular, non-overlapping and contiguous (without holes) areas called tiles.

If blocks are used, as defined in subclause 5.2.1.7, the block containing the first component of the top-left pixel of the frame shall begin at the first byte of the sample data. Otherwise (blocks are not used), the most significant bits of the first component of the top-left pixel of the frame shall be located at the most significant bits of the first byte of the sample data.

The size in bytes of the associated media sample or item shall be at least the size in bytes required to store all the components values documented by the uncompressed video configuration as defined in subclause 5.2.

NOTE The sample data can be larger than the size in bytes required to store all the components values, typically to store information in the trailing data. How such additional bytes are handled by a file reader is out of scope of this specification.

4.2 Storage in media tracks

Uncompressed video tracks compliant to this specification are video tracks compliant to ISO/IEC 14496-12 that use a video sample entry with `codingname` equal to `'uncv'`, hereafter called uncompressed video sample entry.

The uncompressed video sample entry shall contain one `UncompressedFrameConfigBox` and one `ComponentDefinitionBox`.

The `compressorname` field of an uncompressed video sample entry should be set to all 0 (empty string). The `depth` field of an uncompressed video sample entry shall be ignored and should be set to 0, the bit depth per component being indicated by the `UncompressedFrameConfigBox`.

The handler type associated with the track is usually `'vide'`, `'auxv'` or `'pict'` but derived specifications may introduce new handler types. The `width` and `height` fields of the sample

entry shall document the exact frame dimension, in pixels, of any sample of the video stream that is described by this sample entry. Consequently, if the frame dimension changes within a video track, multiple sample entries shall be used.

The payload of an uncompressed video media sample consists of one uncompressed frame.

Each uncompressed video media sample is a sync sample. The `SyncSampleBox`, `ShadowSyncSampleBox`, `CompositionOffsetBox` and `CompositionToDecodeBox` shall not be present in the track.

Media tracks containing only non-visual components should be marked as not present in the presentation, i.e. `track_in_movie` flag should not be set.

Media tracks containing per-pixel meta-data components describing pixels in another track should use a track reference of type `'cdsc'` to the track they describe.

4.3 Storage in image items

An uncompressed image compliant to this specification is an image item compliant to ISO/IEC 23008-12 with the `item_type 'unci'`.

An uncompressed image shall be associated with:

- an `UncompressedFrameConfigBox` essential item property, i.e., `essential` shall be equal to 1 for an `UncompressedFrameConfigBox` item property associated with an image item of type `'unci'`,
- a `ComponentDefinitionBox` essential item property,
- an `ImageSpatialExtentsProperty` whose `image_width` and `image_height` fields shall document the exact frame dimension, in pixels, of the reconstructed image, i.e. the size of the image before applying any associated transformative properties.

The payload of an uncompressed image consists of one uncompressed frame.

An uncompressed image may be associated with meta-data information using the various tools defined in ISO/IEC 14496-12 or ISO/IEC 23008-12, such as descriptive item properties.

Uncompressed images containing only non-visual components should be marked as hidden items, i.e. have (`flags` & 1) equal to 1 in their `ItemInfoEntry`.

Uncompressed images containing per-pixel meta-data components describing pixels in another item should use an item reference of type `'cdsc'` to the item they describe.

5 Uncompressed frame description

5.1 Component Definition

5.1.1 Definition

Box Type: `'cmpd'`
Container: Video sample entry, `ItemPropertyContainerBox`
Mandatory: Yes, if `codingname` of the sample entry is `'uncv'` or if the `item_type` of the item is `'unci'`

Quantity: One per uncompressed video sample entry or one associated per uncompressed image item

The `ComponentDefinitionBox` is used to document the types of components present in samples or items associated with this box through the sample entry or through item property association:

- for an uncompressed video or uncompressed image, this box shall be present,
- for other coding formats, this box may be present to describe component characteristics of other coding formats.

Components defined in the `ComponentDefinitionBox` are referenced by indexes in various boxes in this specification. Care has to be taken while removing components from an uncompressed video or image to also remove in other boxes any reference to the removed components. There is no requirement that all components defined in the `ComponentDefinitionBox` are referenced by other boxes.

For all boxes referring to components defined in the `ComponentDefinitionBox` , the associated `ComponentDefinitionBox` is defined as:

- for an uncompressed video, the `ComponentDefinitionBox` present in the same sample entry as the referring box,
- for an uncompressed image, the `ComponentDefinitionBox` associated, through `ItemPropertyAssociationBox`, to the same image item as the referring box.

The `component_index` field value defined in boxes using an associated `ComponentDefinitionBox` indicates the index in the list of components, with value 0 indicating the first component listed in the associated `ComponentDefinitionBox`. The `component_index` field value shall be strictly less than the `component_count` field value of the associated `ComponentDefinitionBox`.

A `ComponentDefinitionBox` may describe two or more components with the same type (e.g., two monochrome components representing different portions of the electromagnetic spectrum), and file readers may require additional information to process the pixel data. How such information is provided to the file reader is out of scope of this specification.

Table 1 – Component types

Value	Description
0	Monochrome component
1	Luma component (Y)
2	Chroma component (Cb / U)
3	Chroma component (Cr / V)
4	Red component (R)
5	Green component (G)
6	Blue component(B)
7	Alpha/transparency component (A)
8	Depth component (D)
9	Disparity component (Disp)
10	Palette component (P) The `component_format` value for this component shall be 0.
11	Filter Array component such as Bayer, RGBW, etc. (FA)
12	Padded component (unused bits/bytes)

13-0x7FFF	ISO/IEC reserved
0x8000-0xFFFF	User-defined component type(s)

Component types reflect only a nominal characterization of the pixel data and how that pixel data should be displayed; precise bandpass limits, for example, are not implied and may differ from image to image. For some component types, some other boxes can provide additional information, such as `ColourInformationBox` or `MasteringDisplayColourVolumeBox`. For other component types, the signalling of such information is out of scope of this specification.

NOTE If the bands need to be identified beyond the component type, it is recommended to use URIs in place of the enumerated `component_type` value.

5.1.2 Syntax

```
aligned(8) class ComponentDefinitionBox extends Box('cmpd') {
   unsigned int(16) component_count;
   {
      unsigned int(16) component_type;
      if (component_type>=0x8000) {
         utf8string component_type_uri;
      }
   } [component_count];
}
```

5.1.3 Semantics

`component_count` indicates the number of components described in this box

`component_type` indicates the type of the component, as defined in **Error! Reference source not found.**

`component_type_uri` indicates a URI describing the user-defined component type

5.2 Uncompressed Frame Configuration

5.2.1 Definition

5.2.1.1 Overview

Box Type: `'uncC'`
Container: Video sample entry, `ItemPropertyContainerBox`
Mandatory: Yes, if `codingname` of the sample entry is `'uncv'` or if the `item_type` of the item is `'unci'`
Quantity: One per uncompressed video sample entry or one associated per uncompressed image item

The `UncompressedFrameConfigBox` describes uncompressed frames that are composed of one or more components. RGB or YUV formats are typical examples of such uncompressed frames for which each colour component is a component of the uncompressed frame. Other component types can also be described, such as non-visual information (e.g., disparity,

transparency) or non-visible information (e.g., infra-red). The type of each component is specified by the `component_type` field in the associated `ComponentDefinitionBox`.

Component values may be absolute values or indexes into a colour palette, with adjustable black and white reference levels. Pattern-based sensor data, such as Bayer, can be described through user-defined patterns, together with various sensor information (polarization, non-uniformity correction, broken pixels, etc.).

For each component, this box specifies the numerical format (e.g., unsigned integer, IEEE 754 binary32 floating point) and bit depth though the `component_format` and `component_bit depth_minus_one` fields.

Pixel data may be interleaved per component, per pixel, per row or per tile, as specified by the `interleave_type` field, with byte alignment for each row and tile specified by the fields `row_align_size` and `tile_align_size`. Pixel data may also be grouped together in blocks, typically to respect endianness constraints, as specified by the `block_size`, `block_pad_lsb` and `block_little_endian` fields.

Some formats, such as YUV video, do not always use the same 2D resolution for each component of the frame. This is indicated by the `sampling_type` field.

The `UncompressedFrameConfigBox` may indicate a profile for this description, allowing faster identification of the class of video data used. Profiles are defined in subclause 5.3.

5.2.1.2 Component assignment

Each pixel in a frame is made of one or more components, where each component is assigned a type (e.g., 'Y', 'U', 'V'), as detailed in 5.1.

The order in which components are specified in the `UncompressedFrameConfigBox` indicates the order in which components are placed into the sample data or within blocks, prior to blocking and endian conversion.

Some formats, such as Bayer image data, contain a 2D array of single-component values, with each individual component value assigned to a component type using a fixed pattern. Such formats are described as mono-component data with no subsampling, use a component type of 11 ('FA'). There shall be at most one component with type 11 ('FA') present in the component list. There may be additional components of other types present, for example to associate an alpha component with a Bayer image. If a component with type 11 ('FA') is present, there shall be a `ComponentPatternDefinitionBox` present in the video sample entry or associated to the item to indicate the pattern of component values.

Some formats code colours according to a set of predefined colours, or palette. Such formats are described as single-component with no subsampling, use a component type of 10 ('P'). There shall be at most one component with type 10 ('P') present in the component list. There may be additional components of other types present, for example to associate per-pixel alpha component with a palette image. If a component with type 10 ('P') is present, there shall be an `ComponentPaletteBox` present in the video sample entry or associated to the item to indicate the palette values.

5.2.1.3 Component size and numerical format

The variable `component_bit_depth` for a component is defined as (`component_bit_depth_minus_one` + 1).

For a given component, the binary representation of each value is given by `component_bit_depth` (the size in bits of each component value) and `component_format`.

The possible values for `component_format` field is defined in Table 2.

Table 2 - Component formats

Value	Description
0	Component value is an unsigned integer coded on `component_bit_depth` bits.
1	Component value shall be an IEEE 754 binary float number coded on `component_bit_depth` bits (e.g., if `component_bit_depth` is 16, then the component value is coded as IEEE 754 "binary16"). For this component format, `component_bit_depth` values shall be 16, 32, 64, 128 or 256; other values are forbidden.
2	Component value is a complex number coded on `component_bit_depth` bits, where the first `component_bit_depth`/2 bits shall represent the real part and the next `component_bit_depth`/2 bits shall represent the imaginary part. Each part shall be coded as an IEEE 754 binary float of the size `component_bit_depth`/2. For this component format, `component_bit_depth` values shall be 32, 64, 128 or 256; other values are forbidden.
3 – 255	ISO/IEC reserved for future definition

If `component_align_size` is 0, the component value shall be coded on `component_bit_depth` bits exactly.

Otherwise (`component_align_size` is not 0), the component value shall be coded as a word `WC` of `component_align_size` bytes, starting on a byte boundary. This implies that some padding bits may be present after the previous component value stored; if such padding bits are present, they shall be set to 0. The least significant bit of the component value shall be located at the least significant bit of `WC`. Padding bits, if present, shall be located at the most significant bits and shall be set to 0. If `components_little_endian` is 0, `WC` shall be stored as a big-endian word. Otherwise (`components_little_endian` is 1), `WC` shall be stored as a little-endian word.

NOTE For example, a pixel with 10-bit unaligned (`component_align_size`=0) R, G and B components, followed by a 1-byte aligned 7-bit A component, stored using Pixel Interleaving, would exist in the sample data as 30 consecutive bits containing R, G and B, followed by 2 implied padding bits for byte alignment, followed by a padding bit then followed by the 7-bit A value (bringing the A value aligned on 1 byte), for a total of 5 bytes.

For each component, `component_align_size` shall be either 0 or such that `component_align_size*8` is greater than `component_bit_depth`. If `components_little_endian` is 1, `block_little_endian` shall be 0 and `component_align_size` of each component shall be different from 0.

Storage of aligned component values is illustrated in Figure 27**Error! Reference source not found.**.

5.2.1.4 Tiling

The uncompressed frame data is structured in a 2D grid of tiles, where the number of tiles in the horizontal direction (resp. vertical direction) is specified by the variable `num_tile_cols_minus_one`+1 (resp. `num_tile_rows_minus_one`+1). Tiles allow grouping together the component values of pixels close to each-other (i.e., in the same spatial region of the frame).

All tiles have the same width and height. The frame width (resp. height) shall be a multiple of `num_tile_cols_minus_one`+1 (resp. `num_tile_rows_minus_one`+1). The tile width is `w/(num_tile_cols_minus_one+1)`, with `w` the frame width, and the tile height is `h/(num_tile_rows_minus_one+1)`, with `h` the frame height.

If the width (resp. height) in a source image is not an integer multiple of `num_tile_cols_minus_one`+1 (resp. `num_tile_rows_minus_one`+1), an application creating the tiled frame shall pad the source image with an appropriate number of columns to the right (resp. rows to the bottom), and the original image dimension shall be documented using a `CleanApertureBox` for media tracks or a clean aperture transformative item property for image items. The width and height fields of the sample entry or the `image_width` and `image_height` fields of the `ImageSpatialExtentsProperty` shall specify the padded width and height.

Tiles shall be stored in raster-scan order: the top-left tile is stored first, followed by the tile to its right, and the first tile of a tile row is stored after the last tile of the previous tile row.
Within a tile, component values shall be stored in raster-scan order, left-to-right and top to bottom: for a given component, the value for pixel {x+1,y} is stored after (but possibly not contiguous with) the value for pixel {x,y}, and the value for pixel {x,y+1} is stored after (but possibly not contiguous with) the value for pixel {x, y}.

5.2.1.5 Sampling type

All components in a frame either have the same dimensions or use pre-defined sampling modes, indicated by the `sampling_type` field. Possible values for this field are described in Table 3.

NOTE1 Sampling type values restrict the possible interleaving modes since the number of values is not the same for each component.

NOTE2 Derived specifications may further restrict the usage of sampling types with tiling.

Table 3 – Sampling type values

Value	Definition
0	No subsampling All components have the same width and height as the frame. NOTE The tile width and height are not restricted. Derived specifications may further restrict this, for example to enforce width and height to be multiples of 2 in case Y, U and V components are present
1	YCbCr 4:2:2 subsampling

	This value shall only be set if all three `component_type` values 'Y', 'U' and 'V' are present in the component list and if the tile width is a multiple of 2. If this value is used, the `interleave_type` field shall be set to 0 (Component Interleaving), 2 (Mixed Interleaving) or 5 (Multi-Y Pixel Interleaving). The width and height of the 'Y' component are the width and height of the frame. The height of the 'U' and 'V' component is the same as the height of the 'Y' component. The width of the 'U' and 'V' component is half the width of the 'Y' component. Components of other types may be present, and have the same dimension as the 'Y' component. Pixels {x,y} and {x+1,y}, with x%2==0, share the same component values 'U' and 'V'. If `row_align_size` is not 0 and `interleave_type` is 0: • `row_align_size` shall be a multiple of 2 • the row alignment for components of type 'U' and 'V', as defined in 5.2.1.7, shall be done using `row_align_size/2` If `tile_align_size` is not 0: • `tile_align_size` shall be a multiple of 2 • the tile alignment for components of type 'U' and 'V', as defined in 5.2.1.7, shall be done using `tile_align_size/2`
2	YCbCr 4:2:0 subsampling This value shall only be set if all three `component_type` values 'Y', 'U' and 'V' are present in the component list and if both tile width and height are multiple of 2. If this value is used, the `interleave_type` field shall be set to 0 (Component Interleaving) or 2 (Mixed Interleaving). The width and height of the 'Y' component are the width and height of the frame. The width of the 'U' and 'V' component is half the width of the 'Y' component'. The height of the 'U' and 'V' component is half the height of the 'Y' component. Components of other types may be present, and have the same dimension as the 'Y' component. Pixels {x,y}, {x+1,y}, {x,y+1} and {x+1,y+1} with x%2==0 and y%2==0, share the same component values 'U' and 'V'. If `row_align_size` is not 0 and `interleave_type` is 0: • `row_align_size` shall be a multiple of 2

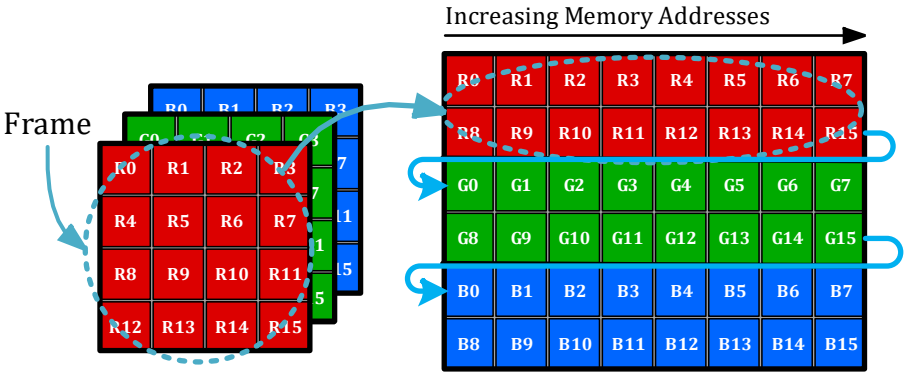
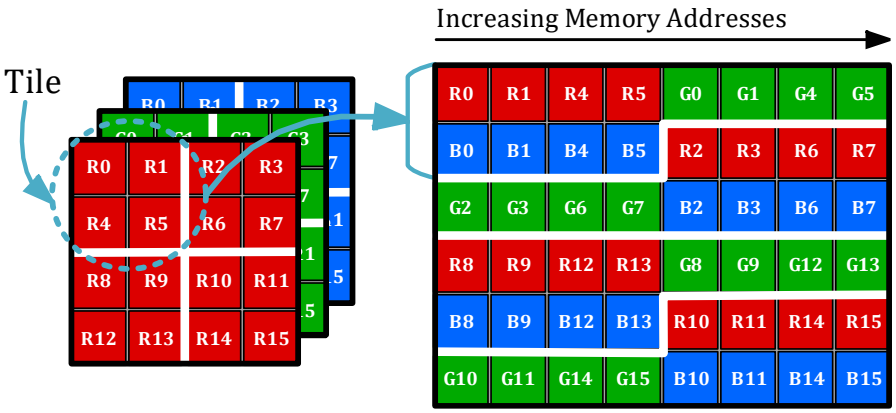
	• the row alignment for components of type 'U' and 'V', as defined in 5.2.1.7, shall be done using `row_align_size/2` If `tile_align_size` is not 0: • `tile_align_size` shall be a multiple of 4 • the tile alignment for components of type 'U' and 'V', as defined in 5.2.1.7, shall be done using `tile_align_size/4`
3	YCbCr 4:1:1 subsampling This value shall only be set if all three `component_type` values 'Y', 'U' and 'V' are present in the component list and if tile width is a multiple of 4. If this value is used, the `interleave_type` field shall be set to 0 (Component Interleaving), 2 (Mixed Interleaving) or 5 (Multi-Y Pixel Interleaving). The width and height of the 'Y' component are the width and height of the frame. The height of the 'U' and 'V' component is the same as the height of the 'Y' component. The width of the 'U' and 'V' component is the width of the 'Y' component divided by 4. Components of other types may be present, and have the same dimension as the 'Y' component. Pixels {x,y} , {x+1,y} , {x+2,y} , {x+3,y}, with x%4==0, share the same component values 'U' and 'V'. If `row_align_size` is not 0 and `interleave_type` is 0: • `row_align_size` shall be a multiple of 4 • the row alignment for components of type 'U' and 'V', as defined in 5.2.1.7, shall be done using `row_align_size/4` If `tile_align_size` is not 0: • `tile_align_size` shall be a multiple of 4 • the tile alignment for components of type 'U' and 'V', as defined in 5.2.1.7, shall be done using `tile_align_size/4`
4-0xFF	reserved

5.2.1.6 Interleaving

The interleaving of pixels within a tile (or within the frame if a single tile is used) is indicated by `interleave_type`, as defined in Table 4. The `interleave_type` shall apply to all listed components unless stated otherwise in Table 4.

NOTE The interleaving describes the layout of values within the sample data. Positioning of sample data within the container file is done according to ISO/IEC 14496-12.

Table 4 – Interleaving types

Value	Definition
0	Component Interleaving For a given component, values for all pixels of a tile shall be located sequentially in the sample data. Component values shall be located in the order the components were declared. This mode can be used with a single tile, as illustrated in Figure 1, or with multiple tiles, as illustrated in Figure 2. In both examples, R is the first component, G is the second and B is the third.  **Figure 1 – Single tile Component Interleaving example**  **Figure 2 – Multiple tiles Component Interleaving example**
1	Pixel Interleaving For a given pixel, the component values shall be located one after the other, in the order the components are declared, with the first component of a pixel following the last component of the previous pixel. Pixel Interleaving is only permitted when all components are at the same resolution. This mode can be used with a single tile, as illustrated in Figure 3, or with multiple tiles, as illustrated in Figure 4.

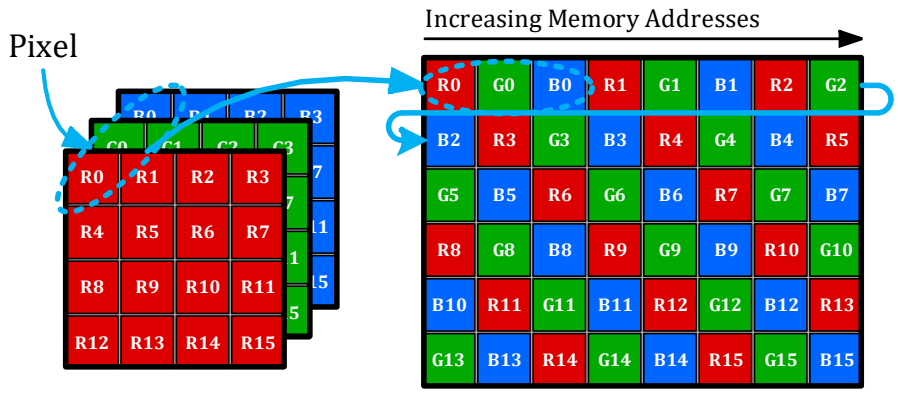


Figure 3 – Single tile Pixel Interleaving example

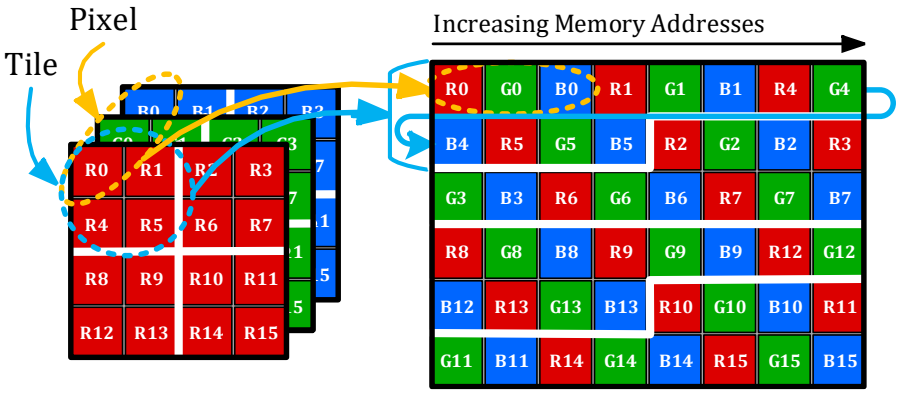


Figure 4 – Multiple tiles Pixel Interleaving example

2	Mixed Interleaving This value shall only be used if `sampling_type` field value is not 0, if `sampling_type` value explicitly allows this interleaving mode and if the 'U' and 'V' components are consecutive in the component list, i.e. 'V' component immediately follows 'U' component or 'U' component immediately follows 'V' component. For all components other than 'U' and 'V', component values for all pixels shall be located as specified for Component Interleaving in the order they are declared. The 'U' and 'V' component values shall be located as specified by Pixel Interleaving; the first value for component 'U', if declared first, or 'V' otherwise, shall be located after the last value of the component preceding 'U' or 'V' if any, or first in the tile otherwise. NOTE This mode is typically used to store YUV 420 data with all Y components followed by interleaved U and V components. Figure 5 illustrates this mode for `sampling_type`=2 and Figure 6 illustrates this mode for `sampling_type`=1.

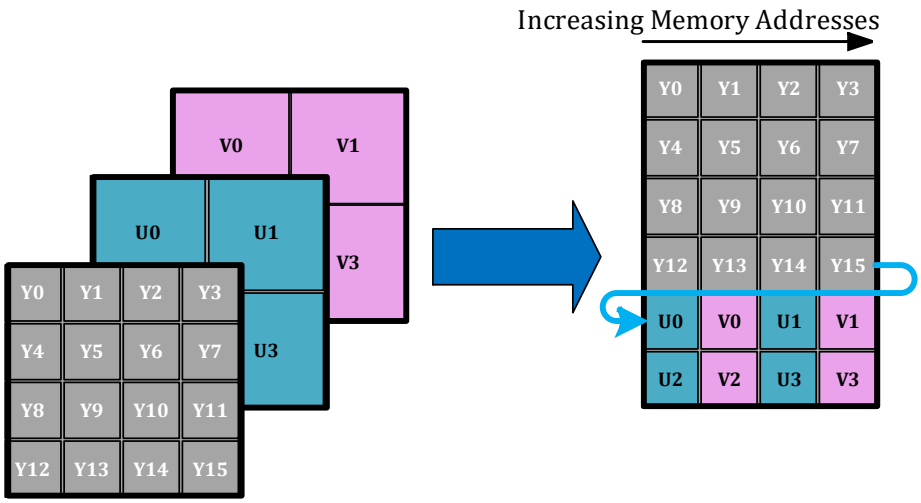


Figure 5 – Mixed Interleaving for YUV 420 example

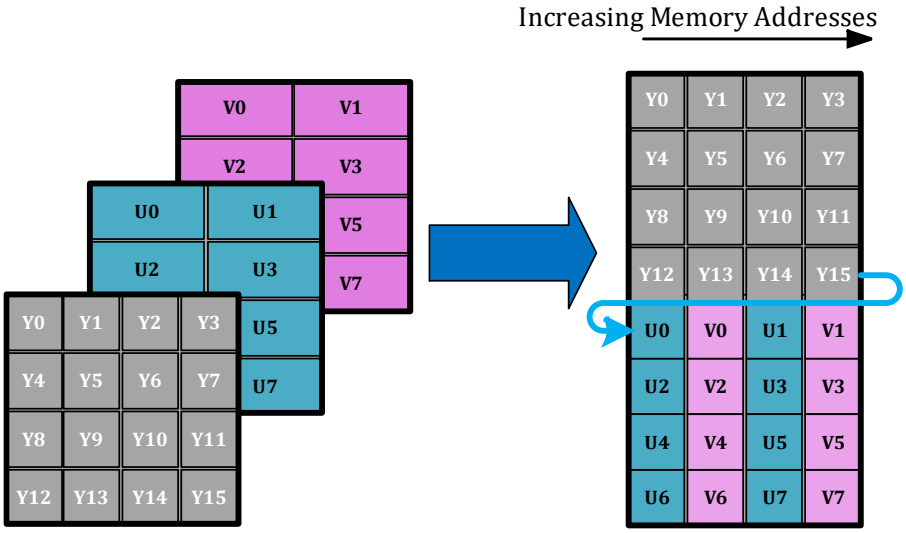


Figure 6 – Mixed Inter leaving for YUV422 example

3	Row Interleaving For a given row, component values for each component shall be located sequentially, in the order the components are declared. Each non-first row in the tile shall be located after the previous row. If multiple tiles are used, the first component of the first row of a tile shall be located after the last component of the last row of the previous tile. Row Interleaving is only permitted when all components are at the same resolution. This mode can be used with a single tile, as illustrated in Figure 7, or with multiple tiles, as illustrated in Figure 8.

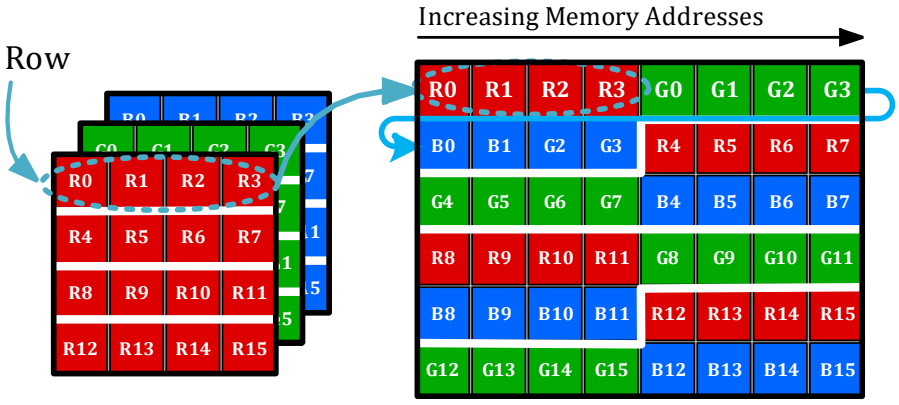


Figure 7 – Single tile Row Interleaving example

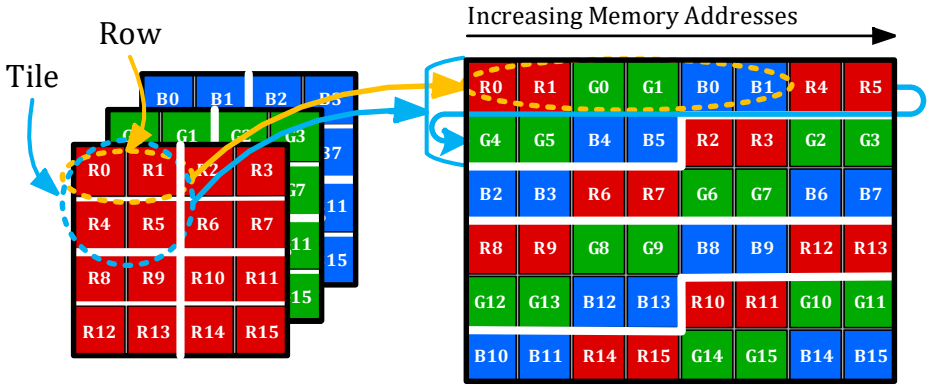


Figure 8 – Multiple tiles Row Interleaving example

4	Tile-Component Interleaving For each component, in the order they are declared, the values for each component for all tiles shall be located sequentially, with values within the tile located per each individual tile. This mode shall not be used if only a single tile is defined. This mode is illustrated in Figure 9.

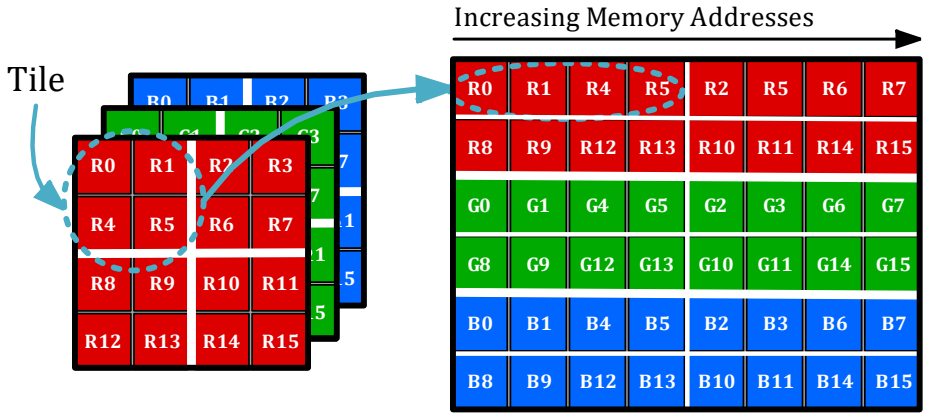


Figure 9 – Tile-Component Interleaving example

5	Multi-Y Pixel Interleaving Component values shall be organized as with `interleave_type`=1 (Pixel Interleaved), and the list of components shall contain multiple

'Y' components representing the multiple 'Y' component values associated with a single pair of 'U' and 'V' component values.
The 'U' and 'V' components shall only be listed once in the list of components.

There are no restrictions on the order of declaration of the 'Y' components; derived specifications may further restrict this.

Regardless of how the multiple 'Y' components are interleaved with the 'U' and 'V' components, the 'Y' component values shall be stored in left-to-right, top-to-bottom order.

If the chrominance components are subsampled according to 4:2:2 subsampling (`sampling_type=1`), the component list describes two horizontally consecutive pixels, and thus shall include two 'Y' components, representing the two 'Y' values associated with a single pair of 'U' and 'V' component values. For example, {'U', 'Y', 'V', 'Y'} for `sampling_type`=1 (YUV 4:2:2) indicates that the Y value of the second pixel shall be located after the 'V' component value.

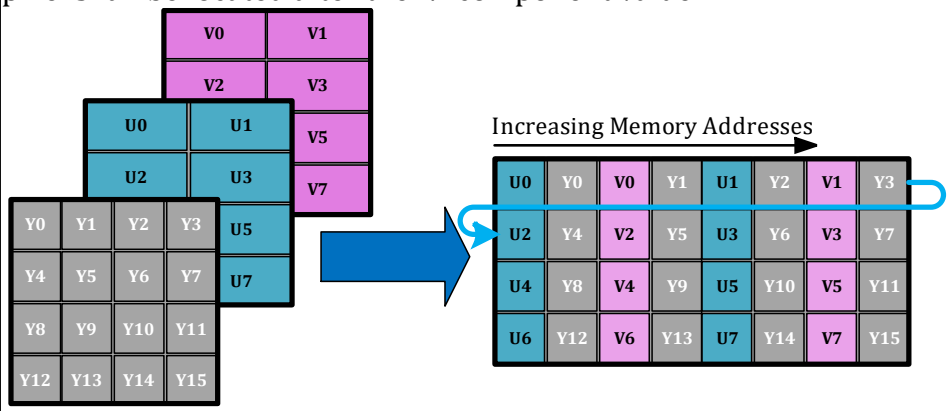


Figure 10 – 'U', 'Y', 'V', 'Y' 4:2:2 example

If the chrominance components are subsampled according to 4:1:1 subsampling, the component list describes four horizontally consecutive pixels, and thus shall include four Y components, representing the four Y values associated with a single pair of U and V component values. For example, {'U', 'Y', 'Y', 'V', 'Y', 'Y'} for `sampling_type=3` (YUV 4:1:1) indicates that the 'Y' value of the second pixel shall be located after the 'Y' value of the first pixel, that the 'Y' value of the third pixel shall be located after the 'V' component value and that the 'Y' value of the fourth pixel shall be located after the 'Y' value of the third pixel.

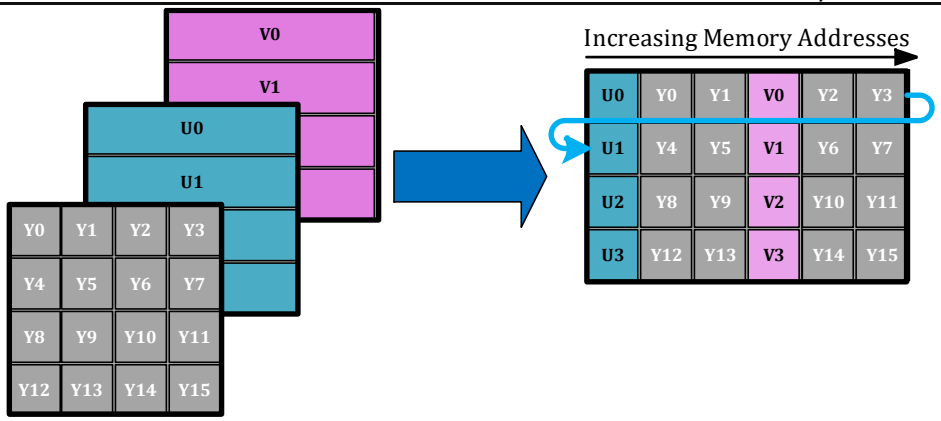


Figure 11 – 'U', 'Y', 'Y', 'V', 'Y', 'Y' 4:1:1 example

This interleave mode shall not be used for other forms of chrominance subsampling.

6-0xFF	Reserved

5.2.1.7 Alignment of values to byte boundaries

Padding bits can be included within the sample data to ensure that component values or pixels are aligned to specific byte boundaries. In addition, the end of rows and tiles can be aligned to specific byte boundaries. This may be to ensure individual component values or pixels are CPU addressable (e.g., not split across two 32-bit words), and rows and tiles of images are aligned to key boundaries (e.g., to memory pages or disk drive sectors).

Component values can be stored either directly in the sample data or inside fixed-size blocks. The block size in bytes is specified by the `block_size` field.

If the block size is 0 (blocking is not used within the sample data):

- no padding is performed between values, except potentially:
  - in-between component values if some components use a `component_align_size` different from 0
  - after the last component value of a pixel if `pixel_size` is set
  - after the last value in a row if `row_align_size` is set
  - after the last value in a tile if `tile_align_size` is set
- `block_pad_lsb`, `block_little_endian` and `block_reversed` shall all be 0,
- values shall be stored most-significant bit first.

Storage of component values with `block_size`=0 is illustrated in Figure 25**Error! Reference source not found.**.

If the block size is not 0:

- each block shall consist in exactly `block_size` bytes and shall contain exactly one or more component values, i.e. the bits representing a component value are always contained within a single block,
- by default, within a block, the first value of the first component of the top-left pixel of the frame shall be stored in the highest order bits of the first block (big-endian "left to right" format). The second value (for next component or next pixel of same component) shall be stored within the next set of bits in the block up until the last component value being stored in the least significant bits of the block,

- if the block size in bits is greater than the size in bits of component values present in the block, padding bits are present. When present, these padding bits shall be located either in the most or in the least significant bits,
- the block size shall be greater than or equal to the maximum value of `component_align_size` for each component.

If `pad_unknown` value is 1, the padding bits values are not restricted, otherwise the padding bits shall all be set to 0.

NOTE1 Padding bits can be more than 8 (one byte). Unrestricted padding bit values is typically used for legacy content with no restrictions on padded values. Derived specification can further restrict usage of `pad_unknown`.

When `block_pad_lsb` is 0, the least significant bit of the last component value in the block shall be located at the least significant bit of the block and, consequently, any padding bits are located in the most significant bits of the block. Otherwise (`block_pad_lsb` is 1), the most significant bit of the first component value in the block shall be located at the most significant bit of the block and, consequently, any padding bits are located in the least significant bits of the block.

The order of storage, in the sample data, of the bytes within the block is determined by the endian storage type. Figure 12 shows a 32-bit block with bits labelled from 0 (least significant bit) to 31 (most significant bit) with the following component assignments:

Component 1: Type 4 (Red) – 8 bits
Component 2: Type 5 (Green) – 8 bits
Component 3: Type 6 (Blue) – 8 bits

Padding bits are located at the least significant bits of the block.
The least significant bit of each component is labelled '0' and the most significant bit '7'.

Figure 12 – Assignment of 8-bit RGB components to a 32-bit block

If `block_little_endian` is 0 (big endian storage), the most significant byte of the block (Component 1 in Figure 12) shall be stored in the lowest storage byte location in the sample data. The remaining bytes shall be stored in increasing address spaces (Component 2 then 3 in Figure 12) and the least significant byte of the block (Padding bits in Figure 12) shall be stored in the highest address location. The big-endian storage arrangement for the component assignments of Figure 12 is shown in Figure 13.

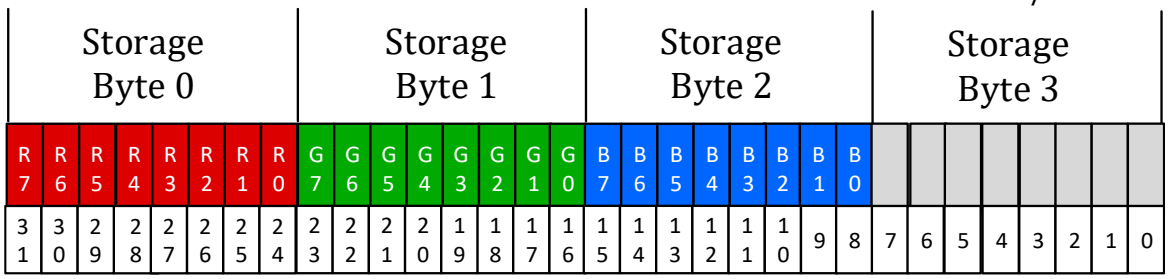


Figure 13 – Big-endian storage of the block defined in Figure 12

If `block_little_endian` is 1 (little endian storage), the least significant byte of the block (Padding bits in Figure 12) shall be stored in the lowest storage byte location. The remaining bytes shall be stored in increasing address space (Component 3 then 2 in Figure 12) and the most significant byte of the block (Component 1 in Figure 12) shall be stored in the highest address location. The little-endian storage arrangement for the block in Figure 12 is shown in Figure 14.

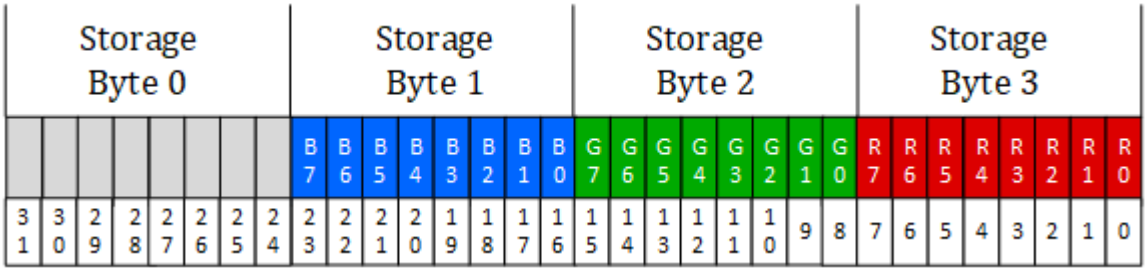


Figure 14 – Little-endian storage of the block defined in Figure 12

If `block_reversed` is 0, the components order within the block shall be the component declaration order. Otherwise (`block_reversed` is 1), the partition of components into blocks shall remain unchanged, but the components order within each block shall be inverted, without changing the location of padding bits. `block_reversed` shall be 0 if `block_little_endian` is 0.

NOTE2 Reversing the order of components in a block is typically used with little-endian storage when multiple blocks are needed to store a single pixel value, to ensure the low bits of a block are either a continuation of the previous pixel or a start of a new one.

An example usage of `block_reversed`=1 is shown in Figure 23 and Figure 24. Figure 29 illustrates usage of `block_reversed`=1 with multi-Y interleave mode.

If `pixel_size` is 0, no additional padding is present after each pixel. Otherwise, let `PixelBytes` be the number of bytes required to contain all component values of a single pixel, including any padding resulting from `block_size` or `component_align_size`. `pixel_size` shall be greater than or equal to `PixelBytes`, and the number of padding bytes present after the last byte of each pixel shall be `pixel_size` - `PixelBytes`. Usage of `pixel_size` different from 0 is illustrated in Figure 28.

NOTE3 `pixel_size` is typically used to ensure the start of every pixel falls on a significant architectural boundary (e.g., all pixels start on 8-byte boundaries).

`pixel_size` shall be 0 if `interleave_type` is different from 1 or 5.

If both `block_size` and `pixel_size` are non-zero, then the first component value of any pixel shall be the first value stored in the first block for this pixel. This implies that the last block of

any pixel may contain additional padding bits. For example, for a 10-bit RGBA image stored using `block_size=4` and `pixel_size=10`, the first block for each pixel will contain the R, G, and B components and 2 bits of padding while the second block for each pixel will only contain the A component value and 22 bits of padding, and two additional bytes of padding will be present after the second block.

NOTE4 Specifying `pixel_size` together with `block_size` ensures no block contains values from different pixels.

When `block_size` and `pixel_size` are both non-zero, the additional padding at the end of each pixel is not required to be a multiple of the block size.

Rows of tiles shall be byte-aligned at the end of the row:

- if the block size is 0, padding bits until byte alignment may be present after the last component value of the last pixel of each row,
- Otherwise, the last block of each row (containing the last component value of the last pixel of each row) may contain more padding bits than previous blocks.

If `row_align_size` is 0, no additional padding is present at the end of rows of tiles. Otherwise, let `RowSize` be the number of bytes required to contain, for a given row `R`:

- all values of all components of row `R` if `interleave_type` is 1 or 5 or if `interleave_type` is 2 and component type is 'U' or 'V' (including all component, block and pixel padding within and at the end of the sample data for row `R`),
- all values of the current component for row `R` (including all component and block padding within the sample data for row `R`) otherwise.

The number of padding bytes present after the last byte of the current row shall be the smallest integer `N`, with `N`>=0, such that `(RowSize+N) % row_align_size` is 0.

If `tile_align_size` is 0, no additional padding is present at the end of tiles. Otherwise, let `TileSize` be the number of bytes required to contain, for a given tile `T`:

- all values of the current component for all rows of the tile `T` (including all component, block and row padding within and at the end the tile `T`) if `interleave_type` is 4,
- all values of all components for all rows of the tile `T` (including all component, block, pixel and row padding within and at the end of the sample data for the tile `T`) otherwise.

The number of padding bytes present after the last byte of the last row of the tile shall be the smallest integer `N`, with `N`>=0, such that `(TileSize+N) % tile_align_size` is 0.

`tile_align_size` shall be 0 if a single tile is used.

NOTE5 Block alignment allows groups of consecutive component values, regardless of interleaving type, to be aligned within an easily addressable value; e.g., multiple components can be packed into 32-bit values but no component value spans two consecutive 32-bit values. For example, a 10-bit RGBA image using pixel interleaving with no component padding and 32-bit (4-byte) block padding on the right will produce a repeated pattern of four 32-bit blocks describing 3 pixels [{R,G,B}, {A, R, G}, {B, A, R}, {G, B, A}]. If tile (resp. row) padding is also specified to be 32-bit (4-byte) then the start the next tile (resp. row) will start at the next 4-byte boundary, regardless of whether space for the next component value was available in the previous block. For example, with 4-byte tile alignment, if the width of the tile is not a multiple of 3, the last block of the line will only contain 1 or 2 component values but all 4-bytes of the last block will be present in the sample data.

5.2.2 Syntax

```
aligned(8) class UncompressedFrameConfigBox extends FullBox('uncC', 0, 0)
{
   unsigned int(32) profile;
   unsigned int(16) component_count;
   {
      unsigned int(16) component_index;
      unsigned int(8) component_bit_depth_minus_one;
      unsigned int(8) component_format;
      unsigned int(8) component_align_size;
   } [component_count];
   unsigned int(8) sampling_type;
   unsigned int(8) interleave_type;
   unsigned int(8) block_size;
   bit(1) components_little_endian;
   bit(1) block_pad_lsb;
   bit(1) block_little_endian;
   bit(1) block_reversed;
   bit(1) pad_unknown;
   bit(3) reserved = 0;
   unsigned int(8) pixel_size;
   unsigned int(32) row_align_size;
   unsigned int(32) tile_align_size;
   unsigned int(32) num_tile_cols_minus_one;
   unsigned int(32) num_tile_rows_minus_one;
}
```

5.2.3 Semantics

`profile` indicates the predefined configuration for this uncompressed frame configuration, as defined in Table 5. Value 0 indicates this uncompressed frame configuration may or may not be compliant to a profile. If not 0, all remaining fields in the box shall have the values indicated in Table 5 for this profile. The four-character code `'gene'` is reserved and shall be interpreted as a value of 0.

`component_count` indicates the number of components present in the frames described by this uncompressed frame configuration

`component_index` indicates the index of the component listed in the associated `ComponentDefinitionBox`

`component_bit_depth_minus_one` indicates the size in bits minus one of the values for the component

`component_format` indicates the format of the component, as defined in Table 2

`component_align_size` indicates the component byte-alignment size, as described in subclause 5.2.1.3

`sampling_type` indicates the sampling mode as defined in Table 3

`interleave_type` indicates the interleaving type of the components, as defined in Table 4

`block_size` indicates the size in bytes of blocks, as described in subclause 5.2.1.7

`components_little_endian` indicates that components are stored as little endian, as described in subclause 5.2.1.3

`block_pad_lsb` indicates padding bits location, as described in subclause 5.2.1.7

`block_little_endian` indicates block endianness, as described in subclause 5.2.1.7

`block_reversed` indicates if component order is reversed within the block, as described in subclause 5.2.1.7

`pad_unknown` indicates that the value of padded bits in blocks or padded components is unknown

`pixel_size` indicates the total number of bytes (including component and block padding) used to store all components for a single pixel when component values for that pixel are interleaved, as described in subclause 5.2.1.7

`row_align_size` indicates the padding between rows, as described in subclause 5.2.1.7

`tile_align_size` indicates the padding between tiles, as described in subclause 5.2.1.7

`num_tile_cols_minus_one` plus one indicates the horizontal number of tiles in the frame

`num_tile_rows_minus_one` plus one indicates the vertical number of tiles in the frame

5.2.4 Examples (Informative)

The following figures illustrate the assignment of three components to 32-bit word blocks and associated storage configurations for endian and block padding options. The least significant bit of the block or of each component is labelled '0', the most significant bit is labelled '31'. The three components for most of the examples are defined as follows:

Component 1: Type 4 (Red) – 9 bits
Component 2: Type 5 (Green) – 10 bits
Component 3: Type 6 (Blue) – 9 bits

Figure 15 shows the assignment of the components to a 32-bit block (`block_size`=4) with padding at the most significant bits (`block_pad_lsb`=0).

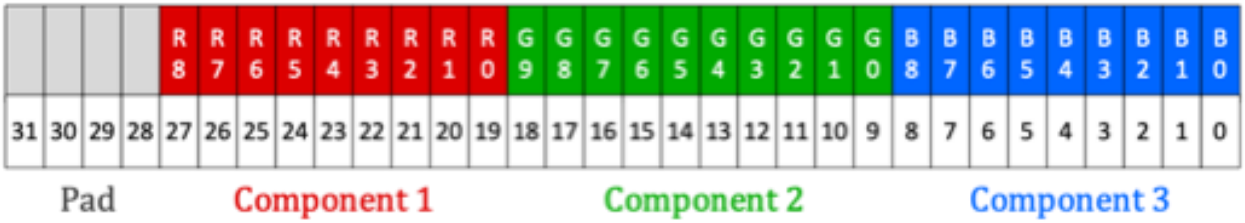


Figure 15 – Example assignment of RGB to 32-bit block, padded at most significant bits

Figure 16 shows the assignment of the components to a 32-bit block (`block_size`=4) with padding at the least significant bits (`block_pad_lsb`=1).

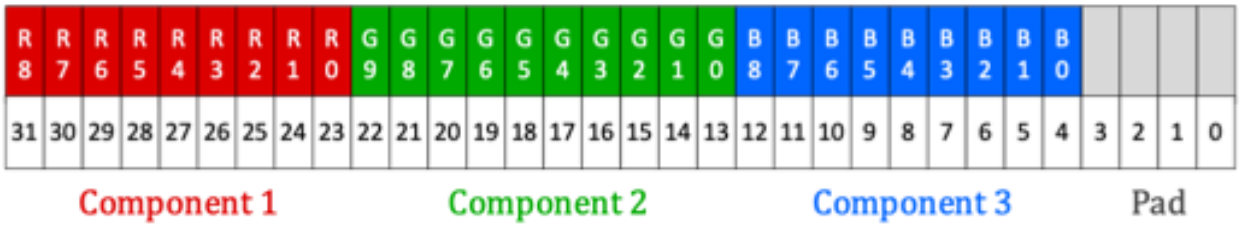


Figure 16 – Example assignment of RGB to 32-bit block, padded at least significant bits

The remaining figures in this section show various combinations of endian, block padding, block reverse, component alignment and pixel alignment options. “Storage Byte 0” (resp. “Storage Byte 3”) in the figures is the first (resp. last) byte stored in the sample data for the illustrated pixel or block.

Figure 17 illustrates a 32-bit block stored in big endian (with `block_little_endian`=0) and padding bits at least significant bits (`block_pad_lsb`=1).

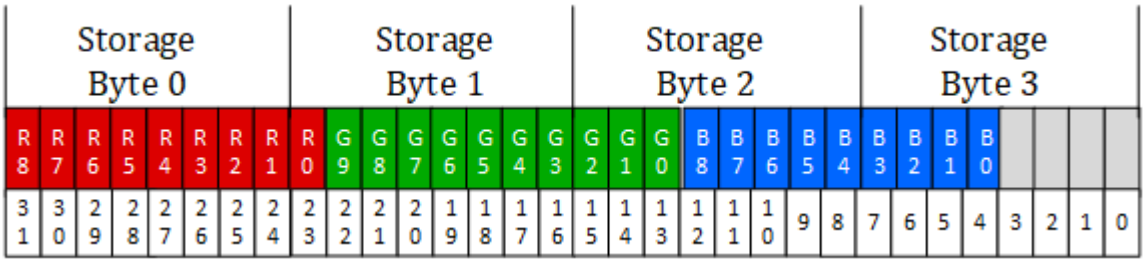


Figure 17 – Big-endian block padded at least significant bits

Figure 18 illustrates a 32-bit block stored in big endian (with `block_little_endian`=0) and padding bits at most significant bits (`block_pad_lsb`=0).

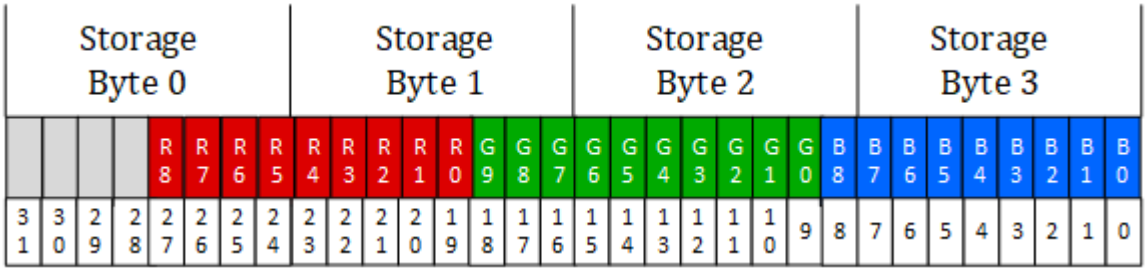


Figure 18 – Big-endian block padded at most significant bits

Figure 19 illustrates a 32-bit block stored in little endian (with `block_little_endian`=1) and padding bits at least significant bits (`block_pad_lsb`=1).

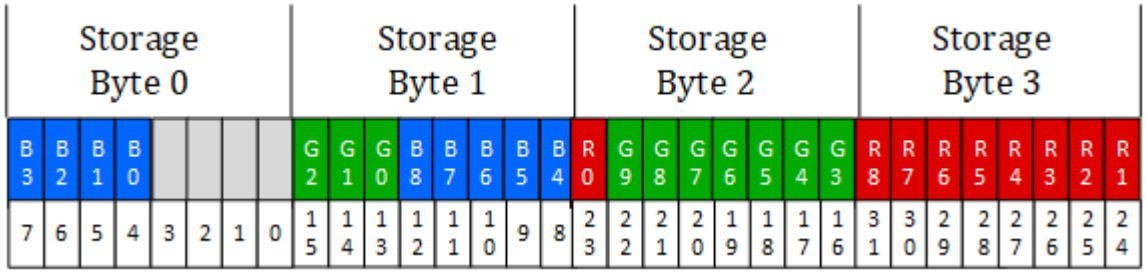


Figure 19 – Little-endian block padded at least significant bit

Figure 20 illustrates Figure 19 with the byte order in the sample data reversed to better show component alignment.

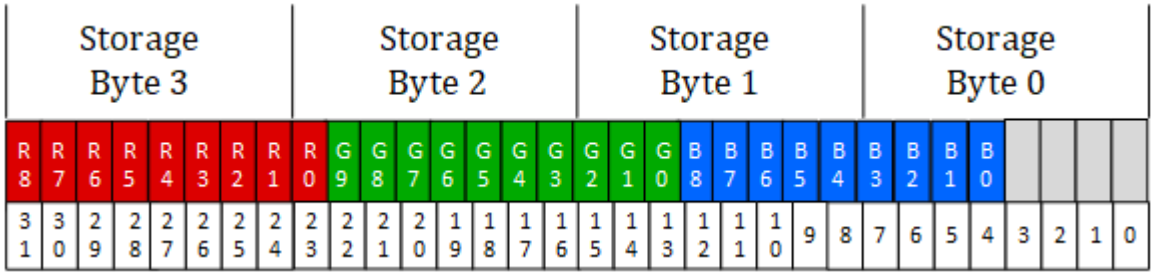


Figure 20 – Little-endian block padded at least significant bit with the storage byte order reversed

Figure 21**Error! Reference source not found.** illustrates a 32-bit block stored in little endian (with `block_little_endian`=1) and padding bits at most significant bits (`block_pad_lsb`=0).

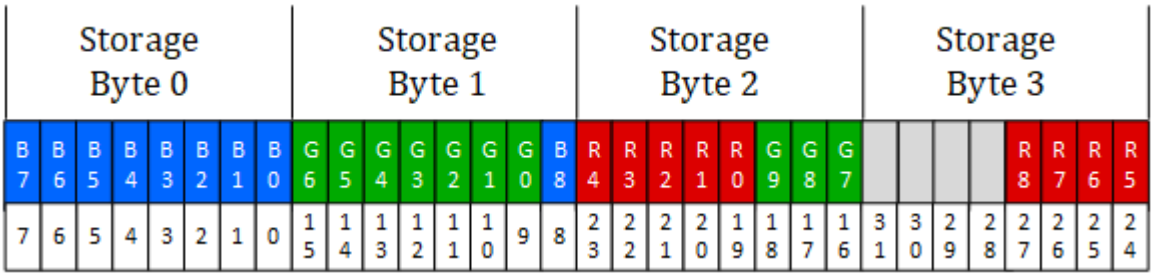


Figure 21 – Little-endian block padded at most significant bit

Figure 22 illustrates Figure 21 with the byte order in the sample data reversed to better show component alignment.

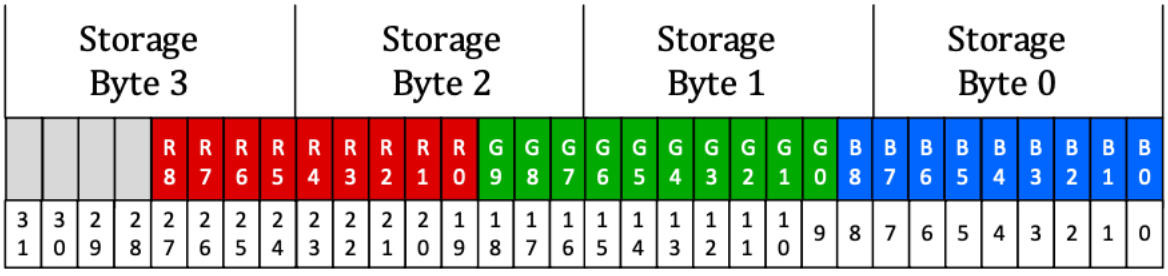


Figure 22 – Little-endian block padded at most significant bit with the storage byte order reversed

Figure 23 illustrates a 32-bit block stored in little endian (with `block_little_endian`=1), padding bits at most significant bits (`block_pad_lsb`=0) and reversed order of component in block (`block_reversed`=1).

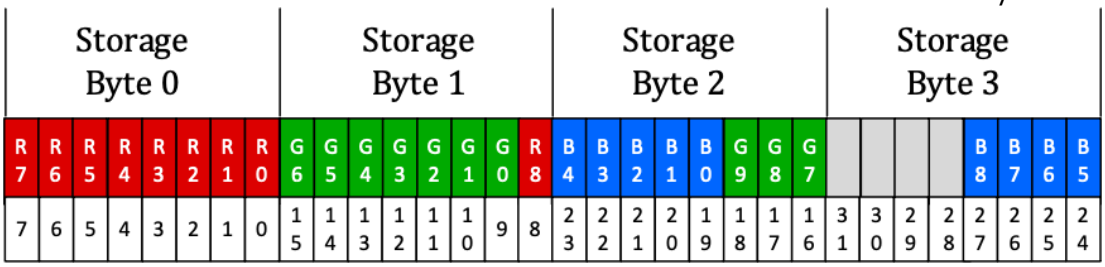


Figure 23 – Little-endian block padded at most significant bit with reversed component order

Figure 24 illustrates Figure 23 with the byte order in the sample data reversed to better show component alignment.

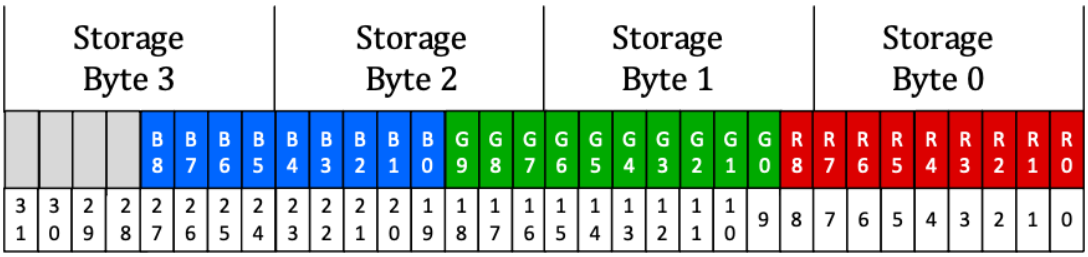


Figure 24 – Little-endian block padded at most significant bit with reversed component order with the storage byte order reversed

Figure 25 shows storage of component Red (9 bits), Green (10 bits) and Blue (9 bits) with no block and no pixel alignment.

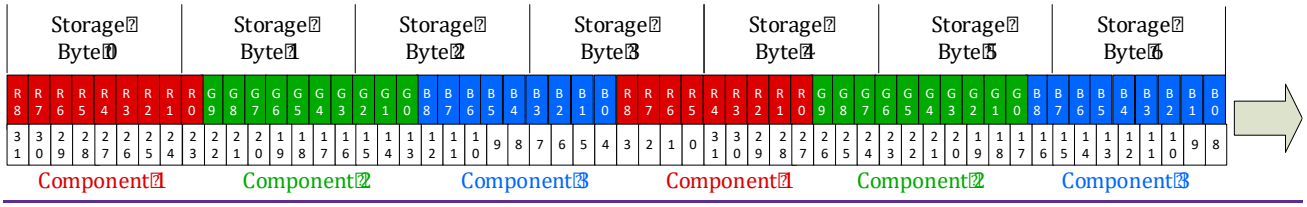


Figure 25 - RGB – red (9 bits), green (10 bits), blue (9 bits), no block, no padding bits

Figure 26 shows storage of components Red, Green, Blue, Alpha, each on 5 bits with no block and 3 bytes per pixels (`pixel_size`=3).

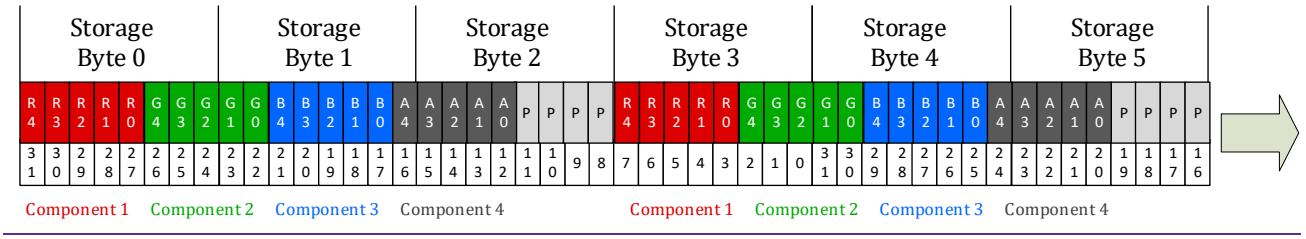


Figure 26 – RGBA (each 5 bits), no block and 3 bytes pixel size

Figure 27 shows storage of components Red (10 bits), Green (12 bits), Blue (10 bits) with no block and `component_align_size`=2 for the Green component only.

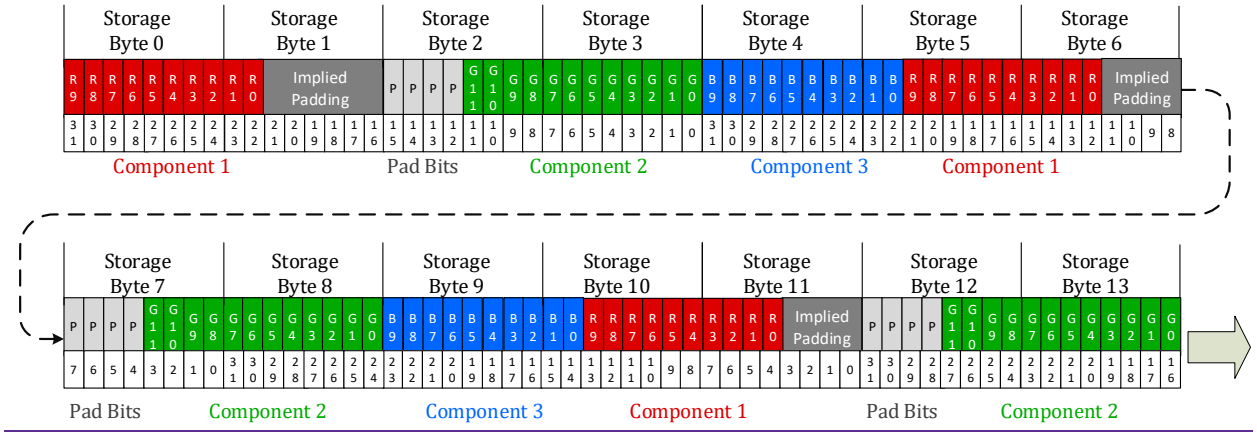


Figure 27 – Red (10-bit unaligned), Green (12-bit, 2-byte alignment), Blue (10-bit unaligned), no block

Figure 28 shows storage of components Red, Green, Blue, Alpha, each on 9 bits within a 32-bit blocks.

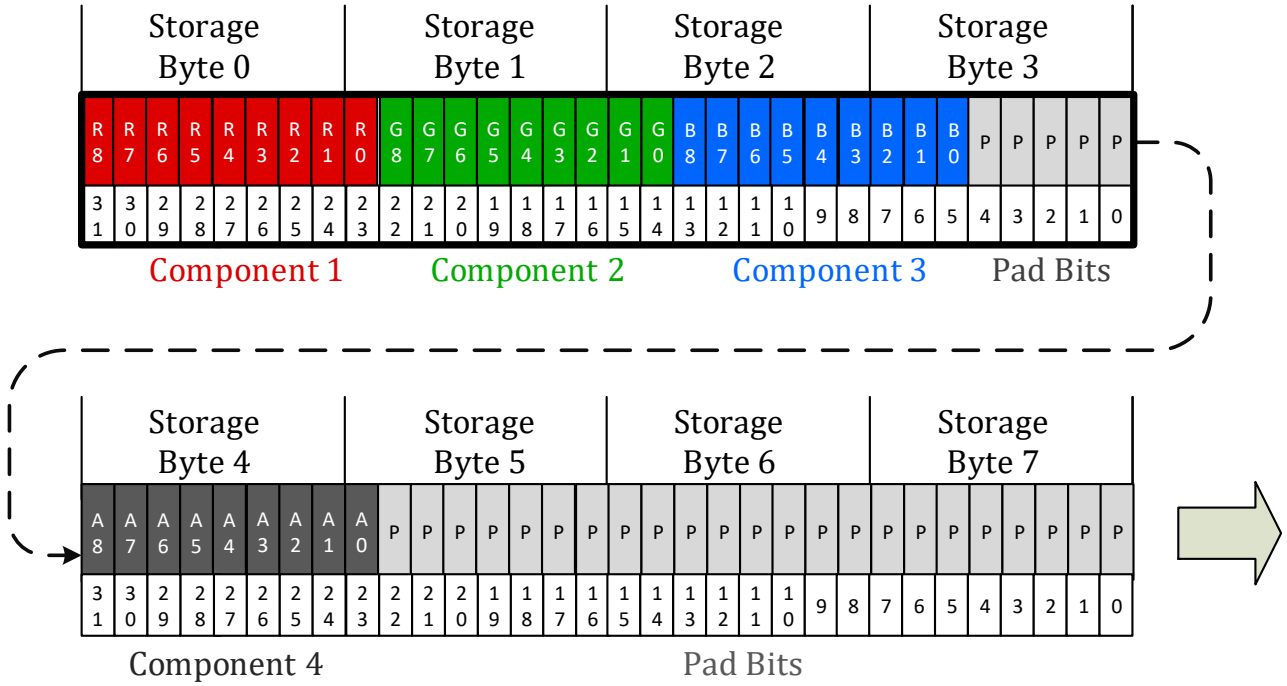


Figure 28 – RGBA (each 9 bits), block size (4 bytes), pixel size (8 bytes)

Figure 29 shows storage of multi-Y interleaved UYVY, each component coded on 10 bits, within 32-bit blocks with `block_little_endian`=1 and `block_reversed`=1, also known as ‘v210’ format.

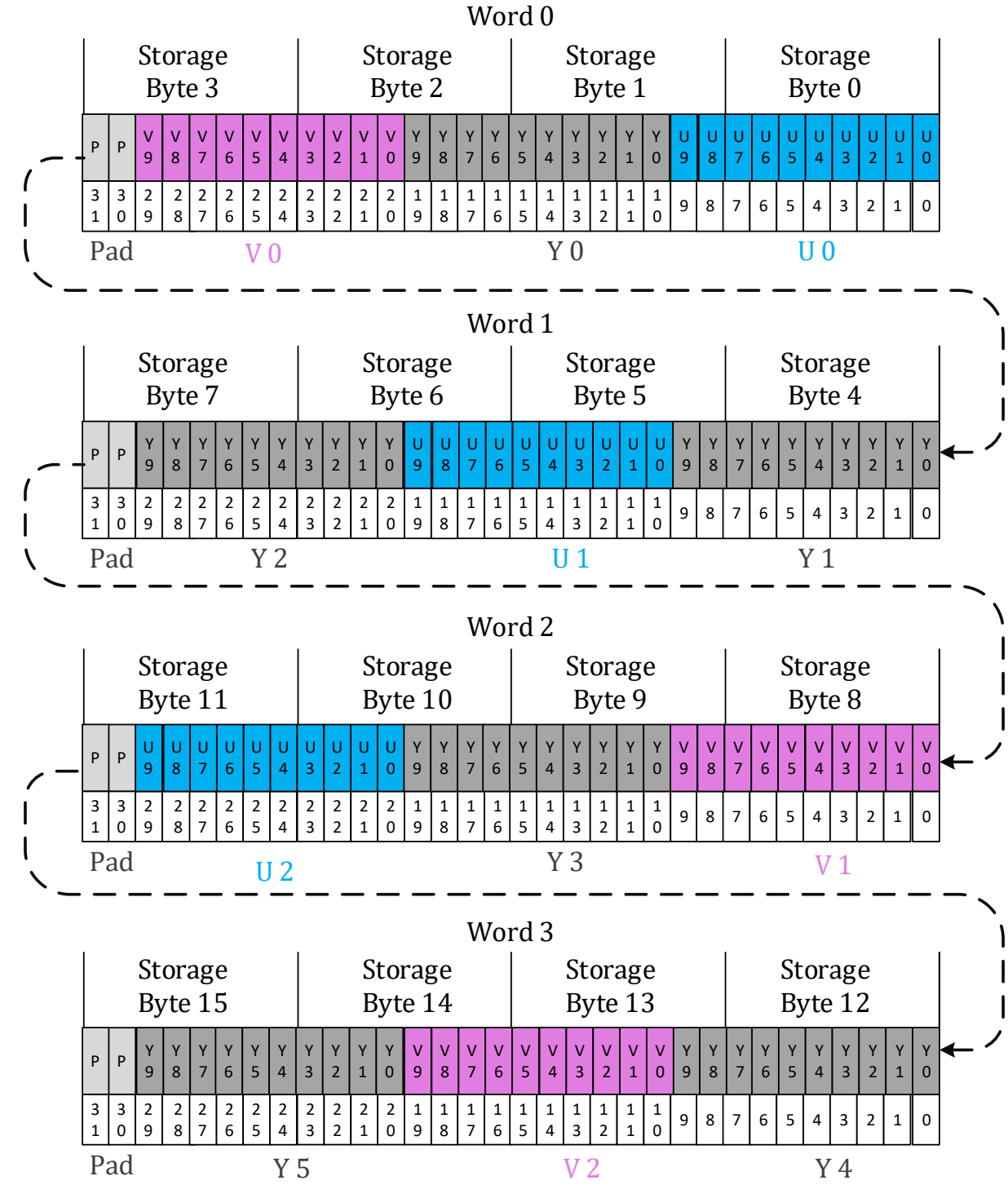


Figure 29 – Multi-Y interleaved component 10-bit UYVY in 32 bit little endian reversed blocks with bytes shown in reverse order per 32-bit word

5.3 Profiles for uncompressed frame configurations

5.3.1 Overview

The four-character codes defined in this subclause identify common formats for which a profile is defined. This profile may be indicated in the `UncompressedFrameConfigBox` to simplify pixel format detection by file processors.

Profiles only provide identification of the pixel format used in uncompressed frames. Image reconstruction might need additional information, such as colour information or chroma location, present in child boxes of the sample entry of the track or associated with the image.

NOTE Some of these profiles are introduced to document existing practices in the industry prior to the definition of this specification.

5.3.2 Predefined configurations

In Table 5 below, the values from the `ComponentDefinitionBox` and the `UncompressedFrameConfigBox` are listed as:

{*profile*, [{*component_type*, *component_bit_depth_minus_1*}], *sampling_type*, *interleave_type*}

The array of {component_type, *component_bit_depth_minus_1*} is given in component order as specified in the `UncompressedFrameConfigBox`, and the number of elements in the array gives the value for `component_count`.

Unless indicated otherwise, all other fields of `UncompressedFrameConfigBox` shall have a value of 0.

Table 5 – Predefined uncompressed frame formats

Profile identifier	Description	Field values for `UncompressedFrameConfigBox`
`'2vuy'`	8 bits YUV 422 packed Cb Y0 Cr Y1	{`'2vuy'`, [{2,7},{1,7},{3,7},{1,7}], 1, 5}
`'yuv2'`	8 bits YUV 422 packed Y0 Cb Y1 Cr	{`'yuv2'`, [{1,7},{2,7},{1,7},{3,7}], 1, 5}
`'yvyu'`	8 bits YUV 422 packed Y0 Cr Y1 Cb	{`'yvyu'`, [{1,7},{3,7},{1,7},{2,7}], 1, 5}
`'vyuy'`	8 bits YUV 422 packed Cr Y0 Cb Y1	{`'vyuy'`, [{3,7},{1,7},{2,7},{1,7}], 1, 5}
`'yuv1'`	8 bits YUV 411 packed Y0 Y1 Cb Y2 Y3 Cr	{`'yuv1'`, [{1,7},{1,7},{2,7},{1,7},{1,7},{3,7}], 3, 5}
`'v308'`	8 bits YUV 444 packed Cr Y Cb	{`'v308'`, [{3,7},{1,7},{2,7}], 0, 1}
`'v408'`	8 bits YUVA 444 packed Cb Y Cr A	{`'v408'`, [{2,7},{1,7},{3,7},{7, 7}], 0, 1}

'y210'	10 bits YUV 422 packed LE Y0 Cb Y1 Cr	{'y210', [{1,9},{2,9},{1,9},{3,9}], 1, 5} block_size shall be 2. block_little_endian shall be 1. block_pad_lsb shall be 1.
'v410'	10 bits YUV 444 packed CbYCr, 2 unused bits	{'v410', [{2,9},{1,9},{3,9}], 0, 1} block_size shall be 4. block_little_endian shall be 1. block_pad_lsb shall be 1. block_reversed shall be 1.
'v210'	YUV 422 10 bits packed CbYCr	{'v210', [{2,9},{1,9},{3,9},{1,9}], 1, 5} block_size shall be 4. block_little_endian shall be 1. block_pad_lsb shall be 0. block_reversed shall be 1. frame width shall be a multiple of 48. row_align_size shall be 0.
'rgb3'	RGB 24 bits packed	{'rgb3', [{4,7},{5,7},{6,7}], 0, 1}
'i420'	YUV 420 8 bits planar YCbCr	{'i420', [{1,7},{2,7},{3,7}], 2, 0}
'nv12'	YUV 420 8 bits semi-planar YCbCr	{'nv12', [{1,7},{2,7},{3,7}], 2, 2}
'nv21'	YUV 420 8 bits semi-planar YCrCb	{'nv21', [{1,7},{3,7},{2,7}], 2, 2}
'rgba'	RGBA 32bits packed	{'rgba', [{4,7},{5,7},{6,7},{7,7}], 0, 1}
'abgr'	RGBA 32bits packed	{'abgr', [{7,7}, {6,7},{5,7},{4,7}], 0, 1}

EDITOR'S NOTE: decide which one we keep, add other configurations, finalize profile names. NB comments on the topic are welcome.

5.4 MIME type sub-parameters

When the 'codecs' parameter of a MIME type is used, as defined in RFC 6381, its value shall be formatted, using field values of the ComponentDefinitionBox and the UncompressedFrameConfigBox, as follows (each element in the value is separated from the previous one using a dot '.'):

- The first element of the value shall be set to 'uncv' for video tracks and to 'unci' for image items,
- If the profile field is not 0, an element equal to the profile four-character code string (case sensitive) is appended
- Otherwise (the profile field value is 0), the string 'gene' (case sensitive) is appended
- Optionally if profile field is not 0, mandatory otherwise:
  - an element equal to the sampling_type expressed in hexadecimal is appended
  - an element equal to the interleave_type expressed in hexadecimal is appended
  - an element equal to the block_size expressed in hexadecimal is appended

- an element equal to '*c*T*r*' is appended, with *c* the number of tile columns in hexadecimal and *r* the number of tile rows in hexadecimal
- for each component, an element equal to '*a*L*b*' is appended, with *a* the hexadecimal value of `component_type` and *b* the hexadecimal value of `component_bit_depth`

Hexadecimal value shall not be prefixed with '0x', and leading zeros should be omitted.

Examples:

- For an uncompressed video with profile `yvyu`: `codecs=uncv.yvyu`
- For an uncompressed video with monochrome and alpha, same sampling (`sampling_type=0`), component interleaving (`interleave_type=0`) with each value stored on 10 bits with 2 bytes blocks little-endian left-padded: `codecs=uncv.gene.0.0.2.1T1.0LA.7LA`
- For an uncompressed video with 2 bits alpha, 10 bits R, B, G, same sampling (`sampling_type=0`), pixel interleaving (`interleave_type=1`) with no specific block alignment (`block_size=0`) and a tiling of 4 horizontal for 3 vertical tiles: `codecs=uncv.gene.0.1.0.4T3.7L2.4LA.5LA.6LA`

6 Component description extensions

6.1 Extensions for uncompressed video and uncompressed images

6.1.1 Overview

The boxes defined in this subclause can be used to provide further information on one or more components present in the uncompressed frame. Presence of these boxes may be mandated by the presence of specific components in the uncompressed frame (e.g., presence of a palette component mandates the presence of a `ComponentPaletteBox`).

Each of the defined boxes can be:

- added to a video sample entry for media tracks
- added as properties associated with an image item.

The syntaxes in this section are given for a video sample entry container and the defined boxes therefore extend `Box` (resp. `FullBox`). When used in an `ItemPropertyContainerBox`, the same syntax applies but the defined boxes extend `ItemProperty` (resp. `ItemFullProperty`).

NOTE Most of the boxes defined in clause 6 may apply to both uncompressed video formats and compressed video formats. Derived specifications may allow usage of these boxes for compressed formats.

6.1.2 Component Palette configuration

6.1.2.1 Definition

Box Type: `'cpal'`
Container: Video sample entry, `ItemPropertyContainerBox`
Mandatory: No
Quantity: Zero or one

The `ComponentPaletteBox` allows describing image data coded through a palette.

This box shall be present if and only if there is an associated `ComponentDefinitionBox` present, in which there is a component defined with a `component_type` value of 10 ('P').

The component value of each pixel gives the index in the palette, and shall be an integer between 0 and `values_count-1`, with value 0 indicating the first entry in the list of pixel values of the palette.

In one `ComponentPaletteBox`, there shall not be any listed component with:

- the same `component_type` as a component listed in the associated `ComponentDefinitionBox`,
- a `component_type` value of 11 ('FA').

6.1.2.2 Syntax

```
aligned(8) class ComponentPaletteBox extends FullBox('cpal', 0, 0) {
   unsigned int(16) palette_components_count;
   {
      unsigned int(16) component_index;
      unsigned int(8) component_bit_depth_minus_one;
      unsigned int(8) component_format;
   } [palette_components_count];

   unsigned int(32) values_count;
   for (i=0; i<values_count; i++) {
      for (j=0; j<palette_components_count; j++) {
         bit(X) component_value;
         aligned(8) bit(0) unused_bits_zero;
      }
   }
}
```

6.1.2.3 Semantics

`palette_components_count` indicates the number of components described by the palette

`component_index` indicates the index of the N[th] component listed in the associated `ComponentDefinitionBox`.

`component_bit_depth_minus_one` indicates the size in bits minus one of values of the N[th] component

`component_format` indicates the format of the N[th] component, as defined in Table 2

`values_count` indicates the number of pixel values following

`component_value` indicates the value of the N[th] component for the given entry. This field is coded on X bits with X = `component_bit_depth_minus_one` + 1 of the N[th] component

`unused_bits_zero` represents the number of unused bits following the component value to achieve byte alignment. These unused bits shall be ignored by the file reader and should be set to 0.

6.1.3 Component Pattern Definition

6.1.3.1 Definition

Box Type: `'cpat'`
Container: Video sample entry, `ItemPropertyContainerBox`
Mandatory: No
Quantity: Zero or one

The `ComponentPatternDefinitionBox` allows describing filter array patterns such as Bayer.

This box shall be present if and only if there is an associated `ComponentDefinitionBox` present, in which there is a component defined with a `component_type` value of 11 ('FA').

The pattern is used to assign the final component type to each value. The pattern is defined at the top-left pixel of the image, and is repeated to cover the entire image. For a pixel position {x, y} in the image, with {0,0} being the top-left pixel, the coded value for that position represents the component type given by position `{width%pattern_width, height%pattern_height}` in the `ComponentPatternDefinitionBox`, and other component values for that pixel, as defined in the `ComponentPatternDefinitionBox`, are not present.

The tile width (resp. height) shall be a multiple of `pattern_width` (resp. `pattern_height`).

In one `ComponentPatternDefinitionBox`, there shall not be a listed component with:

- the same `component_type` as a component listed in the associated `ComponentDefinitionBox`,
- a `component_type` value of 10 ('P').

Examples:

- a BGGR Bayer image will be represented by a single component 11 ('FA') and a pattern with 2 rows, 2 columns indicating components [6,5,5,4],
- A GRBG Bayer image will be represented by a single component 11 ('FA') and a pattern with 2 rows, 2 columns indicating components [5,4,6,5],
- a BGGR Bayer image with an alpha plane following the Bayer data will be represented by two components 11 ('FA') and 7 ('A'), an `interleave_type` of 0 and a pattern with 2 rows, 2 columns indicating components [6,5,5,4],
- a Red-White-Green-Blue sensor image will be represented by a single component 11 ("FA") and a pattern with 4 rows, 4 columns indicating the R, G, B and monochrome components order.

6.1.3.2 Syntax

```
aligned(8) class ComponentPatternDefinitionBox extends FullBox('cpat', 0,
0) {
   unsigned int(16) pattern_width;
   unsigned int(16) pattern_height;
   for (i=0; i< pattern_height; i++) {
      for (j=0; j< pattern_width; j++) {
         unsigned int(16) component_index;
         double(32) component_gain;
      }
   }
}
```

6.1.3.3 Semantics

`pattern_width` indicates the width in pixels of the pattern

`pattern_height` indicates the height in pixels of the pattern

`component_index` indicates the index of the N^{th} component listed in the associated `ComponentDefinitionBox`.

`component_gain` indicates a gain to be applied to the N^{th} component, such as for white balance correction with Bayer imagery, etc. The value shall be coded as an IEEE 754 “binary32”.

NOTE The gain for each component is determined by a user defined method for balancing the levels of all components. The application of the gain terms is also a user defined implementation. This is commonly achieved by setting the gain value to one for the component with the highest signal level and to a value greater than one for the remaining components, thereby balancing all the components. Values saturate at the maximum value of a component.

6.1.4 Component Reference Level

6.1.4.1 Definition

Box Type: `'clev'`
Container: Video sample entry, `ItemPropertyContainerBox`
Mandatory: No
Quantity: Zero or one

The `ComponentReferenceLevelBox` allows describing the minimum and maximum values for components present in the image data.

When this box is present, there shall be an associated `ComponentDefinitionBox` present.

When this box is absent or a component type is not listed in this box, reference white and black values for this component type are derived from the `ColourInformationBox` present in the sample entry of the track or associated with the image item. If the `ColourInformationBox` is absent, the levels for the desired component type are derived as follows:

- For components of type Y, U or V with 8 bits depth, reference black is 16 and reference white is 235
- For components of type Y, U or V with 10 bits depth, reference black is 64 and reference white is for 940
- Otherwise, reference black is 0 and reference white is maximum value for component bit

depth.

When `ComponentReferenceLevelBox` and `ColourInformationBox` are both present and document reference levels for the same component types, information from the `ColourInformationBox` shall be used.

NOTE If the `ColourInformationBox` is present with unspecified `matrix_coefficients` and has `full_range_flag` set to 1, full range is assumed.

Reference levels shall be ignored for non-integer component types.

If `clip_range` is set to 1, `black_level` (resp. `white_level`) indicates the minimum (resp. maximum value) for the component; readers shall clip any value less than `black_level` (resp. greater than `white_level`) to `black_level` (resp. `white_level`).

If `clip_range` is set to 0, readers shall transform the value *N* coded on *k* bits (as read from the sample data) to the value `black_level+`*N*`*(white_level-black_level)/(2`^{k}`-1)` before display or interpretation.

6.1.4.2 Syntax

```
aligned(8) class ComponentReferenceLevelBox extends FullBox('clev', 0, 0)
{
   unsigned int(16) level_count;

   {
      unsigned int(16) component_index;
      unsigned int(1) clip_range;
      bits(7) reserved = 0;
      signed int(32) black_level;
      signed int(32) white_level;
   } [level_count];
}
```

6.1.4.3 Semantics

`level_count` indicates the number of components for which levels are described
`component_index` indicates the index of the Nth component listed in the associated `ComponentDefinitionBox`.
`clip_range` indicates if the levels indicate a clip range or an affine transformation of the Nth component values
`black_level` indicates the black level for the Nth component
`white_level` indicates the white level for the Nth component; this value shall be greater than the `black_level` value.

6.1.5 Polarization Pattern Definition

6.1.5.1 Definition

Box Type: `'splz'`
Container: Video sample entry, `ItemPropertyContainerBox`

Mandatory: No
Quantity: Zero or more

The `PolarizationPatternDefinitionBox` allows describing a filter array pattern for sensors that have polarization patterns implemented on the image sensor. If the sensor also includes a colour or spectral filter, the pattern dimension is not required to match the pattern dimension given in the `ComponentPatternDefinitionBox`.

When this box is present, there shall be an associated `ComponentDefinitionBox` present.

The pattern is used to assign polarization values at the pixel level. The pattern is defined at the top-left pixel of the image, and is repeated to cover the entire image. For a pixel position {x, y} in the image, with {0,0} being the top-left pixel, the polarization value for each pixel is given by position `{width%pattern_width, height%pattern_height}` in the `PolarizationPatternDefinitionBox`.

Polarization angles are specified counter-clockwise, with value 0 indicating an orientation to the right, as illustrated in Figure 30.

The tile width (resp. height) shall be a multiple of `pattern_width` (resp. `pattern_height`).

NOTE If both `pattern_width` and `pattern_height` are 1, this implies that a single polarization angle is given for all pixels in the image.

Multiple `PolarizationPatternDefinitionBox` can be specified if components of the image have different polarization filters. In this case, there shall be at most one `PolarizationPatternDefinitionBox` describing one given component of the image.

Figure 30 illustrates an RGGB Bayer filter array pattern superimposed on a 4x4 pixel grid (filter array pattern with values of 'FA1' through 'FA16') with a 2x2 polarization pattern with angles of 90, 45, 135, and 0 degrees.

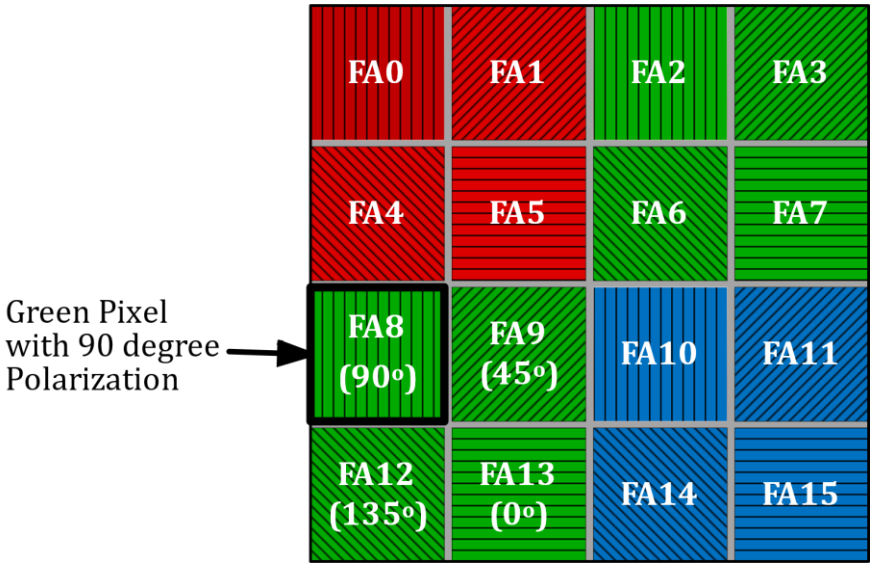


Figure 30 – Polarization example with a filter array pattern component

6.1.5.2 Syntax

```
aligned(8) class PolarizationPatternDefinitionBox extends FullBox('splz',
0, 0) {
   unsigned int(16) component_count;
   {
      unsigned int(16) component_index;
   } [component_count]
   unsigned int(16) pattern_width;
   unsigned int(16) pattern_height;
   for (i=0; i< pattern_height; i++) {
      for (j=0; j< pattern_width; j++) {
         double(32) polarization_angle;
      }
   }
}
```

6.1.5.3 Semantics

`component_count` indicates the number of components to which the polarization pattern applies. If this value is 0, the polarization pattern applies to all components of the image.

`component_index` indicates the index of the component listed in the associated `ComponentDefinitionBox`.

`pattern_width` indicates the width in pixels of the pattern

`pattern_height` indicates the height in pixels of the pattern

`polarization_angle` indicates the polarization angle of the pixel at the defined pattern location. Value shall either be 0xFFFFFFFF, indicating that no polarization filter is present, or shall be an IEEE 754 "binary32" number, indicating the polarization angle in degrees, with a value greater than or equal to 0.0 and strictly less than 360.0.

6.1.6 Sensor Non-Uniformity Correction

6.1.6.1 Definition

Box Type: `'snuc'`
Container: Video sample entry, `ItemPropertyContainerBox`
Mandatory: No
Quantity: Zero or more

The `SensorNonUniformityCorrectionBox` allows describing pixel specific gain and offset corrections.

When this box is present, there shall be an associated `ComponentDefinitionBox` present.

Whether the correction has applied or not to the values stored in the sample data is indicated by the `nuc_is_applied` field.

The Non-Uniform Correction (NUC) `nuc_gain` and `nuc_offset` attributes provide non-uniformity corrections for a sensor. This is commonly used with radiometric imagery. The correction equation is linear: $y =$ `nuc_gain` $* x +$ `nuc_offset`, where x is an input component value as stored in the file. The minimum value of y saturates at a value of 0. In situations where the same component value range is to be maintained after application of the

NUCs, the maximum value of *y* saturates at a level defined by the number of bits in the component value, for instance, 65535 for a 16-bit component.

The NUC coefficients are assigned per pixel through the full image. The NUC gain and offset tables are aligned with the entire image, resulting in a table width equal to `image_width` and table height equal to `image_height`. For a pixel position {x, y} in the image, with {0,0} being the top-left pixel, the NUC coefficients for each pixel are given by position `{x,y}` in the `SensorNonUniformityCorrectionBox`.

Multiple `SensorNonUniformityCorrectionBox` can be specified if components of the image have different gains and offsets. In this case, there shall be at most one `SensorNonUniformityCorrectionBox` describing one given component of the image.

6.1.6.2 Syntax

```
aligned(8) class SensorNonUniformityCorrectionBox extends FullBox('snuc',
0, 0) {
   unsigned int(16) component_count;
   {
      unsigned int(16) component_index;
   } [component_count]
   unsigned int(1) nuc_is_applied;
   unsigned int(7) reserved=0;
   unsigned int(32) image_width;
   unsigned int(32) image_height;
   for (i=0; i< image_height; i++) {
      for (j=0; j< image_width; j++) {
         double(32) nuc_gain;
      }
   }
   for (i=0; i< image_height; i++) {
      for (j=0; j< image_width; j++) {
         double(32) nuc_offset;
         }
      }
   }
}
```

6.1.6.3 Semantics

`component_count` indicates the number of components to which the non uniformity correction pattern applies. If this value is 0, the non uniformity correction applies to all components of the image.

`component_index` indicates the index of the component listed in the associated `ComponentDefinitionBox`.

`nuc_is_applied` if 1, indicates that the corrections have been applied to the component values in the sample data. If value is 0, the component values are still raw, without NUC corrections applied.

`image_width` indicates the width in pixels of the NUC tables, shall be equal to image width

`image_height` indicates the height in pixels of the NUC tables, shall be equal to image height

`nuc_gain` specifies the Non-Uniform Correction (NUC) gain. The value shall be coded as an IEEE 754 “binary32”

`nuc_offset` specifies the Non-Uniform Correction (NUC) offset. The value shall be coded as an IEEE 754 “binary32”

6.1.7 Sensor Bad Pixels Map

6.1.7.1 Definition

Box Type: `'sbpm'`
Container: Video sample entry, `ItemPropertyContainerBox`
Mandatory: No
Quantity: Zero or more

The `SensorBadPixelsMapBox` allows identifying bad pixels on a sensor, i.e. pixels for which at least one component value is corrupted.

When this box is present, there shall be an associated `ComponentDefinitionBox` present.

The box allows entire sensor rows, entire sensor columns, and individual bad pixels to be marked as bad. Pixels that are part of bad rows or bad columns may be listed as part of the list of individually bad pixels, but those pixels are not required to be listed twice.

NOTE1 For example, a sensor system might detect a small region of pixels that show a spurious response where that region intersects with a bad row. The system may report all of the pixels in that region as part of the list of independently bad pixels; it is not required to omit the pixels in the bad row from that list.

NOTE2 If the source image has been padded with rows and/or columns for tile storage, these padded row or columns do not need to be present in the bad pixel map.

Multiple `SensorBadPixelsMapBox` can be specified if components of the image have different sensor defects. In this case, there shall be at most one `SensorBadPixelsMapBox` describing one given component of the image.

6.1.7.2 Syntax

```
aligned(8) class SensorBadPixelsMapBox extends FullBox('sbpm', 0, 0) {
   unsigned int(16) component_count;
   {
      unsigned int(16) component_index;
   } [component_count]
   unsigned int(1) correction_applied;
   unsigned int(7) reserved=0;
   unsigned int(32) num_bad_rows;
   unsigned int(32) num_bad_cols;
   unsigned int(32) num_bad_pixels;
   for (i=0; i< num_bad_rows; i++) {
      unsigned int(32) bad_row;
   }
```

```
   for (i=0; i< num_bad_cols; i++) {
      unsigned int(32) bad_column;
   }
   for (i=0; i< num_bad_pixels; i++) {
      unsigned int(32) bad_pixel_row;
      unsigned int(32) bad_pixel_column;
   }
}
```

6.1.7.3 Semantics

`component_count` indicates the number of components to which the bad pixel map applies. If this value is 0, the bad pixels described apply to all components of the image.

`component_index` indicates the 0-based index of the component listed in the associated `ComponentDefinitionBox`.

`correction_applied` if 1, indicates that component values for bad pixels in the sample data have been corrected; if 0, indicates that no correction has been applied. The method used to apply correction is out of scope of this specification.

`num_bad_rows` the number of full rows in the image that are bad.

`num_bad_columns` the number of full columns in the image that are bad.

`num_bad_pixels` the number of individual bad pixels. This does not include bad pixels that are identified as being part of an entire bad row or bad column.

`bad_row` a row number for which all pixels are bad. Value 0 corresponds to the top row. The value shall be strictly less than image height.

`bad_column` a column number for which all pixels are bad. Value 0 corresponds to the left column. The value shall be strictly less than image width.

`row` the row index of the coordinate pair identifying the location of a bad pixel. Value 0 corresponds to the top row. The value shall be strictly less than image height.

`column` the column index of a coordinate pair identifying the location of a bad pixel. Value 0 corresponds to the left column. The value shall be strictly less than image width.

6.1.8 Chroma Location

6.1.8.1 Definition

Box Type: `'cloc'`
Container: Video sample entry, `ItemPropertyContainerBox`
Mandatory: No
Quantity: Zero or one per video sample entry or associated per item

The `ChromaLocationBox` may be used to describe the chroma subsampling location method used for frames using a YUV subsampling other than YUV444. It shall not be present in the sample entry or associated with the image item for other subsampling types or component formats.

For YUV 4:2:2 or YUV 4:1:1 subsampling, the indicated `Chroma420SampleLocType` as defined in ISO/IEC 23091-2 shall have `VerticalOffsetC` equal to 0.

If `ChromaLocationBox` is not present for frames using a YUV subsampling other than YUV444, the associated `Chroma420SampleLocType` is 0.

If ChromaLocationBox is used as an item property, it shall be marked as essential.

6.1.8.2 Syntax

```
aligned(8) class ChromaLocationBox extends FullBox('cloc', 0, 0) {
   unsigned int(8) chroma_location;
}
```

6.1.8.3 Semantics

chroma_location indicates a Chroma420SampleLocType value as defined in ISO/IEC 23091-2, or the value 6. Value 6 is defined, following Chroma420SampleLocType semantics as defined in ISO/IEC 23091-2, with: VerticalOffsetC=0, HorizontalOffsetC=0 for Cr and HorizontalOffsetC=1 for Cb (ex: DV PAL).

6.1.9 Frame Packing Information

6.1.9.1 Definition

Box Type: 'fpac'
Container: Video sample entry, ItemPropertyContainerBox
Mandatory: No
Quantity: Zero or one per video sample entry or associated per item

The FramePackingInformationBox can be used to describe how two pictures are packed into a frame for stereoscopic imagery.
If not present, each frame consists in a single picture.

If FramePackingInformationBox is used as an item property, it shall be marked as essential.

6.1.9.2 Syntax

```
class FramePackingInfoBox extends FullBox('fpac', 0, 0) {
   unsigned int(4) video_frame_packing;
   unsigned int(4) PackedContentInterpretationType
   unsigned int(1) QuincunxSamplingFlag;
   unsigned int(7) reserved = 0;
}
```

6.1.9.3 Semantics

video_frame_packing, QuincunxSamplingFlag, PackedContentInterpretationType have the same semantics as defined in ISO/IEC 23091-2.

6.1.10 Disparity Information

6.1.10.1 Definition

Box Type: 'disi'
Container: Video sample entry, ItemPropertyContainerBox
Mandatory: No
Quantity: Zero or one per video sample entry or associated per item

The DisparityInformationBox can be used to describe how values of a disparity map should be interpreted. If not present, the mapping from disparity value to distance in meters

uses the default values (`parallax_zero=`2^{N-1}`, parallax_scale=256, dref=300, wref=100`) as defined in ISO/IEC 23002-3.

If this box is used as an item property, it shall be marked as essential.

6.1.10.2 Syntax

```
class DisparityInformationBox extends FullBox('disi', 0, 0) {
   unsigned int(16) component_count;
   {
      unsigned int(16) component_index;
   } [component_count]
   unsigned int(16) parallax_zero;
   unsigned int(16) parallax_scale;
   unsigned int(16) dref;
   unsigned int(16) wref;
}
```

6.1.10.3 Semantics

`component_count` indicates the number of components to which the disparity information applies. If this value is 0, the disparity information applies to all disparity components of the image.
`component_index` indicates the 0-based index of the component listed in the associated `ComponentDefinitionBox`.
`parallax_zero`, `parallax_zero`, `parallax_zero`, `parallax_zero`: have the same semantics as defined in ISO/IEC 23002-3.

6.1.11 Depth Mapping Information

6.1.11.1 Definition

Box Type: `'depi'`
Container: Video sample entry, `ItemPropertyContainerBox`
Mandatory: No
Quantity: Zero or one per video sample entry or associated per item

The `DepthMappingInformationBox` may be used to describe how values in a depth map are transformed into distance values. If not present, the mapping from depth value to distance values follow the default values (`nknear=128, nkfar=128`) as defined in ISO/IEC 23002-3.

If this box is used as an item property, it shall be marked as essential.

6.1.11.2 Syntax

```
class DepthInfoBox extends FullBox('depi', 0, 0) {
   unsigned int(16) component_count;
   {
      unsigned int(16) component_index;
   } [component_count]
   unsigned int(8) nknear;
   unsigned int(8) nkfar;
}
```

6.1.11.3 Semantics

component_count indicates the number of components to which the depth mapping information. If this value is 0, the depth mapping information applies to all disparity components of the image.
component_index indicates the 0-based index of the component listed in the associated ComponentDefinitionBox.
nknear, nkfar near and far distances have the same semantics as defined in ISO/IEC 23002-3.

6.2 Sample group descriptions

6.2.1 Field Interlace Type

6.2.1.1 Definition

The FieldInterlaceType sample group description allows describing the field layout in a sample data in case of interlaced video content. This sample group description is identified by the 'ilce' grouping_type.

If no FieldInterlaceType sample group description is associated to a sample, whether because the sample group description is not present or because the sample is not mapped to any sample group description of this type, the sample data is assumed to be a progressive frame.

6.2.1.2 Syntax

```
aligned(8) class FieldInterlaceDescription {
   unsigned int(2) interlace_type;
   unsigned int(1) interleaved;
   unsigned int(1) first_field_type;
   unsigned int(1) bottom_first;
   bits(3) reserved=0;
}

class FieldInterlaceType extends VisualSampleGroupEntry ('ilce') {
   FieldInterlaceDescription interlace_description;
}
```

6.2.1.3 Semantics

interlace_type indicates the interlacing type. The following values are defined:

- 0: progressive frame
- 1: the sample data contains both top and bottom fields. The duration of each field is half the duration of the associated sample.
- 2: the sample data contains a single field. The video sample entry height shall be twice the number of lines in the field. The duration of the field is the duration of the associated sample.
- 3: the sample data contains both top and bottom fields, but the fields are co-incident in time (sample data is a progressive segmented frame).

interleaved if set to 1, indicates that lines of each field are interleaved in the sample data, i.e. first line of first field followed by first line of second field followed by second line of first field etc.). Otherwise (interleaved set to 0), the first line of the second field follows the last line of the first field in the sample data. This value shall be 0 for interlace_type values other than 1 and 3.

first_field_type if set to 1, indicates that the first field appearing in the sample data is the bottom field. If a single field is present in the sample data, then this field is the bottom field. This value shall be 0 if interlace_type value is 0.

bottom_first if set to 1, indicates that the first field to present is the bottom field. This value shall be 0 if interlace_type value is not 1.

EDITOR's NOTE: We could also use the reserved bits to indicate source origin/copy types for content resulting from telecine (2:2, 3:2 pulldown) if needed. NB Comments and contributions on the topic are welcome.

6.3 Image Item properties

6.3.1 Field Interlace Property

6.3.1.1 Definition

Box Type: `'ilcp'`
Container: `ItemPropertyContainerBox`
Mandatory: No
Quantity: Zero or one per item

The `FieldInterlaceProperty` allows describing the field layout of an image item in case of interlaced video content. In not present, the image item is assumed to be a progressive frame. If present, the property shall be marked as essential.

> NOTE Rendering of images consisting in two non time-coincident fields or in a single field is implementation specific. Derived specification may further constraint usage or rendering of interleaved image items.

6.3.1.2 Syntax

```
class FieldInterlaceProperty extends ItemProperty ('ilcp') {
   FieldInterlaceDescription interlace_description;
}
```

6.3.1.3 Semantics

The semantics are identical to the `FieldInterlaceType` sample group description semantics, using the image item width (resp. height) instead of the video sample entry width (resp. height).

7 Multiple track and items storage

7.1 Overview

Multiple track (resp items) storage allows describing uncompressed video (resp. image) streams whose components are in multiple tracks (resp. items). This process may be used in various cases:

- simplify editing, by modifying only one component (e.g. alpha or depth) without modifying the other components (colour)
- provide players a way to play a subpart of the components (e.g. colour) instead of failing playback due to unsupported configuration
- enable efficient disk and network usage for partial access (spatial, temporal) of a subset of the components. A typical use case is to store all or a subset of samples for the first component followed by samples of the second component etc. by using ISOBMFF sample and item data layout tools.
- use components with different spatial resolutions

7.2 Component video track group

Video tracks containing components of the same media source may be grouped using a `TrackGroupBox` with the `track_group_type` value of `'scvg'`. The pair of `track_group_id` and `track_group_type` identifies a track group within the file. The tracks that contain a particular `TrackGroupTypeBox` having the same value of `track_group_id` and `track_group_type` belong to the same track group.

Tracks belonging to such a group may have different resolutions and different sample entry types, but shall have:

- The same aspect ratio,
- The same temporal layout, i.e. there shall not be any presentation time for which no sample is defined for one or more tracks of the group while a sample is defined for other tracks of the group.

NOTE The component video track group is mainly intended for storage of uncompressed video with components in different tracks, but can be used for a mix of compressed and uncompressed video (e.g. colour compressed, alpha / depth uncompressed). Derived specification may further restrict the kind of components allowed in `'scvg'` track groups, the timescale used, the sample entry format, the visual dimension of the tracks, etc. The display or processing of (part of) a track group with `track_group_type` value of `'scvg'` is application specific.

7.3 Image tiling using ISOBMFF tracks and items

Source image data may be spatially split in different rectangular images, or tiles, each resulting tile being stored in its own track or image item rather than storing the entire image data as a single track or item.

A typical use case is to allow a temporal interleaving of tiles different from what can be achieved with tiling as defined in 5.2.1.4, for which each sample data shall contain all tiles of the associated frame.

Video tracks containing tiles of the same media source may be grouped using a `TrackGroupBox` with the `track_group_type` value of `'stvg'`.

The pair of `track_group_id` and `track_group_type` identifies a track group within the file. The tracks that contain a particular `TrackGroupTypeBox` having the same value of `track_group_id` and `track_group_type` belong to the same track group, i.e. they contain different tiles of the same source image. Tile layouts shall be indicated using the width, height and matrix of the track header. Tracks belonging to the same track group with type `'stvg'` shall not spatially overlap each other. There may be uncovered area(s) in the reconstructed frame, for example when a tile track is removed from the file. Tracks belonging to the same track group with type `'stvg'` are not required to use the same sample entry type.

For image items, tile layouts can be indicated using the grid derived image item.

NOTE Derived specification may further restrict the combination of tiles within the uncompressed video tracks and multi-track tiling or the presence of holes in the reconstructed frame.